

武器装备备件需求预测

赵建忠 徐廷学 叶文 丛林虎 董琪
李海军 吕晓峰 孟蕾 扬子江 岳奎志 著

电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书着重介绍了武器装备备件的基本概念、需求影响因素、需求预测的程序及武器装备备件保障决策时常用的需求预测方法, 主要包括基于相似分析法、模糊层次综合评判法、灰色评估法的备件品种确定方法, 以及基于仿真、时间序列、灰色模型、支持向量机、故障分析、组合模型的备件需求预测方法。

本书可作为武器装备备件订货决策、储存保管、供应管理人员及装备维修保障人员的参考书和培训教材, 也可作为工业部门研究人员、保障性工程师及高等院校装备综合保障、维修保障等专业人员的参考书。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有, 侵权必究。

图书在版编目(CIP)数据

武器装备备件需求预测 / 赵建忠等著. —北京: 电子工业出版社, 2018.10
ISBN 978-7-121-31753-8

I. ①武… II. ①赵… III. ①武器装备—军需保障—需求—预测 IV. ①E075

中国版本图书馆 CIP 数据核字 (2018) 第 218187 号

策划编辑: 张正梅

责任编辑: 刘小琳

印 刷:

装 订:

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 720×1 000 1/16 印张: 19 字数: 372 千字

版 次: 2018 年 10 月第 1 版

印 次: 2018 年 10 月第 1 次印刷

定 价: 98.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888, 88258888。

质量投诉请发邮件至 zltts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式: (010) 88254757。

前言



在人类社会和自然科学问题中，普遍存在预测预报问题，大至宇宙的未来演化、人类历史发展的趋势，小至销售预测、股市预测、天气预报等。预测是在做决策、规划之前必不可少的重要环节，是科学决策、规划的重要前提。

武器装备是由许多部（组）件组成的，在使用期需要对这些部（组）件进行修理和更换。为了缩短武器装备的修理间歇时间，增加可正常工作的时间，对那些易磨损、易故障或因形状复杂、要求高、加工周期长而可能有故障的部（组）件，储备合理数量的备件，以备随时更换，是提高武器装备可用度和降低寿命周期费用，保证能按时发挥战斗效能的有效措施。备件是武器装备维修保障资源中不可缺少的组成部分，是部队成系统、成建制形成作战能力和保障能力的物质基础，备件供应直接影响装备的战备完好性和寿命周期费用。现代高技术战争的实践表明，备件保障已成为战争中谋求战场主动权的有效手段，备件保障能力直接影响武器装备的使用及战备完好性，关系到军事目标的实现和部队“两成两力”的建设。目前，我军建设正处于难得的战略机遇期，新型武器装备不断入列服役，我军武器装备结构原理的复杂和保障任务的增加，对备件需求保障工作提出了更多、更高的要求。

在备件保障问题中，备件需求预测与优化配置是推进“精确保障”的重要手段，是实现“保障有力”目标的重要途径，也是备件保障的重点、难点问题。如果不能科学、合理地确定备件的品种和数量，就将造成装备保障经费的浪费，对装备的战备完好性产生非常大的影响，甚至会影响装备的保障能力，进而影响部队的战斗力，影响部队顺利地完成任务，关系到作战的胜负。因此，武器装备备件需求预测与配置优化技术，是实现武器装备备件供应保障由“规模型”向“精益型”转型的重要手段，是保证备件供应规

划“优生”和“优育”的有效途径，也是装备综合保障方向研究的重点和难点之一。

现有的预测方法分别从不同的角度分析与解决问题，但当运用于武器装备备件保障时，其对问题分析和方法选择缺乏系统的研究，往往预测结果精度不高，方法可扩展性不强，实际效果达不到预期要求。笔者曾在海军一线部队从事装备保障工作多年，后来调入院校从事武器装备综合保障方面的教学工作，同时参与了总部和海军下达的备件消耗规律研究、备件消耗标准制定、备件需求辅助决策支持系统开发等多项科研项目，对武器装备备件供应保障工作有较为全面和系统的了解。在装备保障和教学科研工作过程中，笔者深深体会到备件需求预测与配置优化工作必须做到具体问题具体分析、理论与实践相结合。为有助于武器装备备件保障决策与管理工作的顺利实施，笔者参阅有关预测书刊和资料，吸取各行业预测实例的经验和教训，在个人历年的研究攻关、工作实践和思考，以及在学习工作中陆续发表的备件需求预测相关的论文的基础上，形成了本书的主要内容。

本书在系统分析集成的研究框架下，较为系统地、完整地从事理论和应用方面介绍了备件需求预测的理论与方法。第1章概述，主要介绍武器装备备件的基本概念、备件的分类、备件需求影响因素分析和备件需求特点分析等。第2章武器装备备件需求预测与配置优化基础，主要介绍武器装备备件的消耗模式，备件需求预测的基本原则、一般程序和方法，以及备件配置优化指标与方法等。第3章武器装备备件品种确定方法，主要介绍备件确定常用的方法及其特点，包括相似分析法、模糊层次综合评判法和灰色评估法及其应用等。第4章基于寿命分布和仿真的武器装备备件需求预测方法，主要介绍不同寿命分布类型的备件需求预测模型，以及蒙特卡罗仿真、考虑设备停机的备件需求预测模型、Logistic 回归和马尔可夫过程的备件需求预测模型及其应用等。第5章基于时间序列和回归分析的武器装备备件需求预测方法，主要介绍时间序列法、平稳随机时间序列法和回归分析法及其应用等。第6章基于灰色模型的武器装备备件需求预测方法，主要介绍灰色预测的建模过程，以及含有空缺值序列的 GM(1,1)模型、加入影响因子的灰色模型、灰色-马尔可夫模型的建模过程及其应用等。第7章基于人工智能的武器装备备件需求预测方法，主要介绍 BP 神经网络、支持向量机回归和组合相关向量机预测方法及其应用。第8章基于故障分析的武器装备备件需求预测方法，主要介绍基于故障预测的备件需求预测程序，基于信息融合和 IMUGM(1,1)故障预测、基于信息融合和 IMUGM(1,m,w)故障预测，以及基于故障规律的备

件需求预测方法及其应用等。第9章基于组合模型的武器装备备件需求预测方法,主要介绍组合模型的建模过程,基于回归分析和灰色模型、小波变换和GM(1,1)-ARMA模型、Theil不等系数和指数平滑法、GM(1,1)和支持向量机的备件需求预测方法及其应用等。书中所列的数据和曲线,一般基于海军有关单位的保障数据,某些分析结论可能不适合所有的武器装备,但是分析方法具有通用性。

本书第1、3章由徐廷学编写,第2、6、7、8章由赵建忠编写,第4、5章由叶文编写,第9章由丛林虎编写,全书由赵建忠等统稿。

本书在编写过程中,得到了许多同事、专家的鼓励、支持和帮助,在此一并表示感谢。路明、赵建印、邓建球、张旭、刘勇做了大量的审稿工作,并提出宝贵意见。感谢海军航空大学徐廷学教授及其所指导的金天球、王相飞等研究生的帮助,没有他们协助建模计算和进行令我受益匪浅的讨论,本书是难以完成的。此外,对编写中参阅的有关书刊、文献的作者致以敬意和感谢。特别感谢我的夫人魏炜女士对我工作的理解和支持。

由于笔者水平有限,书中难免有错误和不妥之处,真诚期待广大读者批评和指正。

赵建忠

2018年5月

目 录



第 1 章 概述	1
1.1 备件的基本概念	1
1.2 备件的分类	3
1.2.1 常见的备件分类方法	3
1.2.2 低消耗备件	4
1.2.3 基于模糊隶属度的低消耗备件的定义与界定	4
1.3 备件需求影响因素分析	10
1.3.1 备件需求内在影响因素分析	11
1.3.2 备件需求外在影响因素分析	12
1.3.3 影响因素的量化及规范化处理	16
1.3.4 基于粗糙集属性约简的主要影响因素确定	20
1.4 备件需求层次性分析	22
1.4.1 备件层次需求分析	22
1.4.2 装备系统层次需求分析	23
1.4.3 维修保障组织需求分析	23
1.5 备件需求特点分析	23
1.5.1 备件需求特征	23
1.5.2 常消耗备件需求特点	24
1.5.3 低消耗备件需求特点	25
第 2 章 武器装备备件需求预测与配置优化基础	26
2.1 备件消耗模式分析	26

2.2	备件需求预测的基本原则	27
2.3	备件需求预测的一般程序	29
2.4	备件需求预测的方法	31
2.4.1	定性备件需求预测方法	31
2.4.2	定量备件需求预测方法	32
2.4.3	模型复杂性与模型选择	34
第 3 章	武器装备备件品种确定方法	36
3.1	概述	36
3.2	基于相似分析法的备件品种确定方法	42
3.2.1	相似分析法简介	42
3.2.2	示例分析	44
3.3	基于模糊层次综合评判法的备件品种确定方法	46
3.3.1	模糊层次综合评判法简介	46
3.3.2	模型建立	47
3.3.3	示例分析	50
3.4	基于 D-S 理论和模糊综合评判的备件品种确定方法	52
3.4.1	模型建立	52
3.4.2	实现步骤	54
3.4.3	示例分析	57
3.5	基于灰色评估法的备件品种确定方法	61
3.5.1	评价指标体系的建立	61
3.5.2	评价模型的建立	63
3.5.3	示例分析	65
第 4 章	基于寿命分布和仿真的武器装备备件需求预测方法	69
4.1	基于寿命分布的备件需求预测方法	69
4.1.1	模型假设	70
4.1.2	模型建立	70
4.1.3	示例分析	74
4.2	基于蒙特卡罗仿真的备件需求预测方法	77
4.2.1	模型假设	78
4.2.2	参数确定	78

4.2.3	蒙特卡罗方法	79
4.2.4	仿真模型建立	79
4.2.5	示例分析	81
4.3	基于马尔可夫与蒙特卡罗仿真的备件需求预测方法	82
4.3.1	马尔可夫预测模型	82
4.3.2	备件消耗的马尔可夫性分析	84
4.3.3	基于马尔可夫与蒙特卡罗仿真预测的模型建立	85
4.3.4	示例分析	88
4.4	基于 Logistic 回归和马尔可夫过程的备件需求预测仿真模型	89
4.4.1	模型建立	90
4.4.2	需求发生概率预测	91
4.4.3	需求数量预测	93
4.4.4	误差分析	96
4.4.5	示例分析	97
4.5	考虑设备停机的备件需求预测仿真模型	100
4.5.1	基于最大熵原理与概率加权矩的备件寿命分布确定	101
4.5.2	基于保障度的间断工作备件需求预测模型建立	106
4.5.3	仿真思路及流程规划	109
4.5.4	示例分析	111
第 5 章	基于时间序列和回归分析的武器装备备件需求预测方法	115
5.1	时间序列法概述	115
5.2	基于时间序列的备件需求预测方法	116
5.2.1	算数平均预测法	116
5.2.2	移动平均预测法	117
5.2.3	指数平滑预测法	118
5.2.4	示例分析	122
5.3	基于回归分析的备件需求预测方法	126
5.3.1	线性回归概述	126
5.3.2	一元线性回归模型	127
5.3.3	多元线性回归模型	129
5.3.4	示例分析	131

第 6 章 基于灰色模型的武器装备备件需求预测方法	133
6.1 灰色预测方法概述	133
6.1.1 灰色系统简介	133
6.1.2 灰色系统理论基础	134
6.1.3 灰色预测方法	136
6.2 基于一般序列 GM(1,1) 模型的备件需求预测方法	137
6.2.1 模型建立	137
6.2.2 示例分析	141
6.3 基于含有空缺值序列 GM(1,1) 模型的备件需求预测方法	144
6.3.1 模型建立	144
6.3.2 示例分析	145
6.4 基于加入影响因子改进灰色模型的备件需求预测方法	146
6.4.1 改进灰色预测模型的建立	146
6.4.2 示例分析	147
6.5 基于灰色-马尔可夫模型的备件需求预测方法	155
6.5.1 经典灰色-马尔可夫预测模型	155
6.5.2 改进的灰色-马尔可夫预测模型	158
6.5.3 示例分析	162
第 7 章 基于人工智能的武器装备备件需求预测方法	166
7.1 基于 BP 神经网络的备件需求预测方法	166
7.1.1 BP 神经网络基本原理	167
7.1.2 预测模型的建立	169
7.1.3 示例分析	170
7.2 基于支持向量机 (SVM) 的备件需求预测方法	175
7.2.1 支持向量机概述	176
7.2.2 支持向量机预测原理	177
7.2.3 支持向量机预测模型	185
7.2.4 示例分析	188
7.3 基于复杂时间序列的组合相关向量机备件需求预测方法	192
7.3.1 相空间重构	192
7.3.2 小波变换基本理论	193

7.3.3	小波函数的选取	198
7.3.4	基于组合相关向量机的预测原理	199
7.3.5	示例分析	206
第 8 章	基于故障分析的武器装备备件需求预测方法	213
8.1	故障预测的基本原理分析	213
8.2	基于故障预测的备件需求预测系统构建	215
8.2.1	工作流程	216
8.2.2	数据预处理	216
8.2.3	故障预测	219
8.2.4	备件需求确定	220
8.3	基于信息融合和 IMUGM(1,1) 故障预测的备件需求预测方法	220
8.3.1	问题描述	221
8.3.2	IMUGM(1,1) 模型建立	221
8.3.3	基于信息融合和 IMUGM(1,1) 故障预测的备件需求 预测模型构建	224
8.3.4	示例分析	225
8.4	基于信息融合和 IMUGM(1,m,w) 故障预测的备件需求预测方法	230
8.4.1	问题描述	231
8.4.2	IMUGM(1,m,w) 模型的建立	231
8.4.3	基于信息融合和 IMUGM(1,m,w) 故障预测的备件需求 预测模型构建	234
8.4.4	示例分析	234
8.5	基于故障规律的备件需求预测方法	239
8.5.1	备件需求量计算公式	239
8.5.2	示例分析	242
第 9 章	基于组合模型的武器装备备件需求预测方法	244
9.1	组合预测方法概述	244
9.1.1	一般组合预测模型	245
9.1.2	最优加权系数的确定	246
9.1.3	模型的检验	247

9.2 基于回归分析和灰色模型的备件需求预测方法	247
9.2.1 单个预测模型建立	247
9.2.2 最优组合预测模型	248
9.2.3 示例分析	248
9.3 基于小波变换的备件需求预测方法	251
9.3.1 小波变换理论	251
9.3.2 GM(1,1)-ARMA 模型的改进	254
9.3.3 WGM-ARMA 模型的构建	264
9.3.4 示例分析	265
9.4 基于改进 Theil 不等系数的备件需求预测方法	270
9.4.1 单项预测模型的建立	271
9.4.2 基于 Theil 不等系数的 IOWHA 算子的组合预测模型构建	275
9.4.3 基于 Theil 不等系数的 IOWHA 算子的组合预测模型改进	278
9.4.4 示例分析	282
参考文献	287

第 1 章

概 述

唯物辩证法认为，具体问题具体分析。武器装备使用、维修和保障的特殊性，决定了其备件供应必须要与之相适应。武器装备备件需求预测与配置优化应在研制阶段保障性分析的基础上，充分考虑武器装备使用强度、工作环境、维修模式等因素，以获得合理的备件保障方案。

1.1 备件的基本概念

本书有关备件的相关概念在保持装备既有传统的观念上，基本与 GJB 4355—2002《备件供应规划要求》中关于备件的术语相一致，这里仅介绍与本书内容密切相关的术语与定义。

(1) 备件：维修装备及其主要成品所需的元器件、零件、组件或部件等的统称。元器件是用于电子设备中的基本单元制品，如晶体管、集成电路块等；零件是装备中不可再分解的单个制件，是装备构成的基本单位，如螺栓、螺帽等；部件是装备中由若干零件装配的独立的组成部分，如控制器、继电器等。备件包括可修复备件与不可修复备件。

(2) 可修复备件：装备故障或损坏后，采用经济可行的技术手段修理，能恢复其原有功能的备件，也称为备用件。

(3) 不可修复备件：装备故障或损坏后，不能用经济可行的技术手段加以修

复的备件，也称为修理件。

(4) 现场可更换单元 (Liner Replaceable Unit, LRU): 把一个系统看作由可维修的组件构成，为使该系统恢复到正常状态，可以更换这些组件。更换工作通常在基层级或一级维修设施进行，也就是说，在该系统使用的场所进行。这种更换的目的是，在恢复该系统正常状态的过程中最大限度地减少停机时间。

(5) 车间可更换单元 (Shop Replaceable Unit, SRU): 每个 LRU 通常由分组件或模件构成，而为了易于修复该 LRU，可以在第二级（中继级）或第三级（基地级）维修单位（或车间）内更换该分组件或模件。

(6) 修理周转期 (Turn Around Time, TAT): 指某个故障件从使用装备上撤离开始，通过基地级（或中继级）维修，然后返回使用现场作为备件所经历的时间，也称为周转时间。

(7) 失效: 产品丧失实现规定功能能力的事件，可理解为“元器件、零部件、总成、机器、设备、系统等产品丧失规定的功能”。

(8) 失效率: 也称为故障率，是与固有可靠性对应的自然故障的发生概率，它反映了器件本身的设计、制造水平，是预计备件需求率的重要依据。其定义为“工作到某时刻 t 时尚未失效（故障）的产品，在时刻 t 以后的下一个单位时间内发生失效（故障）的概率”，记作 $\lambda(t)$ 。它反映 t 时刻失效的速率，也称为瞬时失效率。

(9) 备件消耗率: 单位时间内的备件消耗量。备件消耗率的预计是解决好备件供应保障问题的关键和核心，它对备件需求量、采购量、分配量、储存量的确定与优化均有直接的影响。

(10) 备件需求量: 在规定时间内，装备维修保障对备件需求的平均数量称为备件需求量。其可以是单台装备的，也可以是某型装备一个单位的需求量。从定义可以看出，备件需求量与一定使用时间相对应。从平均意义上讲，使用时间长备件需求量就大；反之，备件需求量则小。在实际统计与预测中，备件需求量一般对应于一个批量供应周期。备件需求量包括人为因素造成的需求，如丢失、操作失误、维修中的损坏等。

注：除特殊说明外，本书中的备件需求量等同于备件消耗量。

(11) 备件需求率: 在规定条件下，装备在单位时间内对备件需求的平均数量。其表达式为

$$a = \frac{N}{t}$$

式中： t 为备件供应时间，对于导弹装备而言，平时备件供应时间为日历时间；

N 为该段时间内用于维修装备的备件需求量。

1.2 备件的分类

随着武器装备构造日趋复杂,品种和数量繁多、科技含量高成为备件的显著特点。备件分类是备件管理现代化的重要内容,是实现备件自动化管理的基础和前提。

1.2.1 常见的备件分类方法

(1) 按可恢复性分为可修复备件与不可修复备件。

(2) 按使用阶段分类:初始备件,指在装备形成战斗力的初始保障期内,装备使用与维修所需的备件。初始备件是从承制方那里采购(或者是承制方制造,或者是承制方从转承制方、供应方处得到)的。后续备件,指装备已形成初始战斗力且订购方已具有备件保障能力后,在规定时间内装备使用与维修所需补充的备件,该类备件是承制方向使用方推荐的订货备件,作为已消耗掉的初始备件的补充件,由使用方选择订货。

(3) 按技术状况和质量等级分类:新品(一级品),指未经使用并符合技术条件的备件;堪用品(二级品),指使用过或库存超期,经检修后符合技术标准的备件;待修品(三级品),指有故障或技术状况不明,需要检修的备件;报废品(四级品),指符合报废标准并已批准报废的备件。

(4) 按结构属性分类:电子件备件,其寿命分布一般按指数分布处理;机械件备件,其寿命分布一般可按正态分布处理,属于限寿件,按寿命长短又可进一步划分为全寿件(整个服役期内除执行规定的保养外,不需翻修)、单寿件(对应一个浴盆盆底长度即翻修时限)、短寿件(基层级预防维修件,也称易损件备件);其他件备件,如橡塑件、木材等,一般不假定寿命分布,其寿命按经验数据给出,也归入限寿件。

(5) 按供货源分为通用件(市售件)、专用件(制造厂产品)。

(6) 按寿命分布分类:指数寿命件,其故障率为常数;正态寿命件,其故障率依赖时间;威布尔寿命件,其故障率呈现递增趋势。

(7) 按结构特点分类:部附件,包括设备、部件、附件和组合件,一般是可修复备件;零配件,包括零件和元器件,一般是不可修复备件。

(8) 按备件存放时的状态分类: 待修备件, 不能使用的可修复备件; 准备发放备件, 具有使用功能和满足操作规范的可以发放的备件, 该备件可以是新的、修复的或大修好的。

(9) 按使用对象分为部队排故修理备件、工厂大修备件和战场抢修备件。

(10) 按备件所占的资金比例分类: A 类备件, 其品种只占全部库存备件的 5%~15%, 资金占 60%~80%; B 类备件, 其品种占 20%~30%, 资金占 20%~30%; C 类备件, 其品种占 60%~80%, 资金占 5%~15%。

(11) 按维修级别分为基层级备件、中继级备件和基地级备件。

(12) 按购置费用分为昂贵件和一般件。

(13) 按对系统功能的影响分为关键件和一般件。

(14) 按备件的技术特性分类: 标准件, 即按国家标准制造的备件; 专门件, 即按主机制造厂自定的标准系列制造的备件; 特制件, 即非标准的特制备件。

1.2.2 低消耗备件

关于低消耗备件的定义, 国内外备件领域都有相关的研究成果, 但以定性分析为多。总结国内外关于低消耗备件定义的研究, 主要有以下四种观点:

- (1) 认为低消耗备件为年消耗数不超过 2 件的备件;
- (2) 认为低消耗备件为年申请数低于 6 件的备件;
- (3) 认为低消耗备件为 180 天需要 1 件或近 1 年需要 2 件的备件;
- (4) 认为低消耗备件为过去的 1 年里需求量小于 4 件的备件。

以上四种观点均明确给出了“低消耗”量的标准, 为理论研究带来了方便。但上述任何一种定义仅仅代表了一种主观认识, 并没有对认识的科学性进行论证, 所以从理论上讲, 这种缺乏论证的主观认识无法作为进一步研究的依据。在实际工作中, 各单位对备件的使用管理现状也不尽相同, 对低消耗备件的定义无法取得一致, 认识的不统一, 也给低消耗备件的管理带来了不便。

因此, 上述对低消耗备件定义的研究虽然得出了各自的结论, 但无法将其推广应用于实际工作中, 可操作性较差。

1.2.3 基于模糊隶属度的低消耗备件的定义与界定

在备件管理研究中, 研究者对低消耗备件的认识不尽一致。例如, 有研究认为低消耗备件为过去一年消耗量在 2 件或 2 件以下的备件; 也有研究认为低消耗

备件为年申请数小于6件的备件。这种定义存在不足之处：首先，这种定义过于“精确”，“低”是用自然语言定义的一个模糊的概念，这种模糊性是难以量化的，如果一定要为一个模糊的概念下一个确定性的定义，那么得到的结论未必能很好地反映所讨论现象的本质。其次，这种定义主观性太强，难免会脱离客观实际，可操作性较差。在实际工作中，各单位应该根据本单位的实际来确定“低消耗”的范围。因为“低”的程度是一个不确定的因素，而这种程度的定义又对库存结构的确定有着举足轻重的影响，所以应该由也只能由决策者自身做出判断。最后，对低消耗备件的定义不能仅仅考虑备件某一年的消耗数据，因为备件的需求一般是连续的，历史数据与未来的备件需求具有很大的关联性，因此在判断某项备件是否属于低消耗备件时，仅仅凭借一年的数据是没有说服力的。

鉴于以上原因，为有效管理低消耗备件，本书将“低消耗备件”看作一个动态的、相对的概念，用模糊数学原理确定“低消耗备件”这一模糊集的隶属函数，依据隶属函数对低消耗备件进行定义和分类。

1. 模糊集与隶属度

众所周知，每个概念都有一定的内涵和外延，它们是描述概念的两个方面。外延是指概念反映的那一类对象，内涵是指概念反映的客观对象的本质。用集合的观点来看，一个概念的内涵就是集合的定义，外延则是组成该集合的所有元素。普通集合描述概念的特点：一个对象要么符合该概念，要么不符合该概念，二者必居其一，没有模棱两可的情况。而对于模糊概念来说，一个对象是否符合它，不能简单地用“是”或“否”来回答，因为对象既不完全符合也不完全不符合，而是在某种程度上符合该概念。

普通集合 A 可用由它到集合 $\{0,1\}$ 上的映射 $A(u)$ 来描述，如果不再用它二值逻辑的规定，把特征函数值的范围从集合 $\{0,1\}$ 扩充到在区间 $[0,1]$ 上连续取值，那么一个对象符合某概念的程度可以用区间 $[0,1]$ 内的数来表示，对象所对应的数值越靠近1，表示该对象符合概念的程度越大，反之越小。这样一来，普通集合 A 就相应地扩充为一个带有不分明边界的模糊集，从而模糊概念也就可用这样的模糊集来表达。

定义 1.1 设在论域 U 上给定了映射：

$$u_{\bar{A}}:U \rightarrow [0,1], u \rightarrow u_{\bar{A}}(u)$$

则称 u 确定了 U 上的一个模糊子集，记为 \bar{A} 。 $u_{\bar{A}}$ 称为模糊子集 \bar{A} 的隶属函数， $u_{\bar{A}}$ 在 $u \in U$ 点处的值 $u_{\bar{A}}(u)$ 称为 u 对 \bar{A} 的隶属度，它表示 u 属于 \bar{A} 的程度。为方便起

见,通常将模糊子集简称为模糊集,且把 $u_{\bar{A}}$ 与 $u_{\bar{A}}(u)$ 均简记为 $\bar{A}(u)$ 。模糊集 \bar{A} 完全由其隶属函数所描述,即只要给定隶属函数 $\bar{A}(u)$,模糊集 \bar{A} 也就完全确定,因此,考查某一模糊概念的定义,重点是确定这一模糊概念的隶属函数。

模糊度是描述模糊集模糊程度的数量指标,用符号 $d(\bar{A})$ 表示。不同的模糊集其模糊程度是不一样的,即使是同一论域上的模糊集,它们的模糊程度也可能有较大区别。普通集是不模糊的,标志它的模糊程度的数值应该为零;而对 $\forall u \in U$,若 $\bar{A}(u) \equiv 0.5$,此时 \bar{A} 最为模糊,标志其模糊程度的数值应为1。此外,标志模糊集 \bar{A} 的模糊性大小的量还应该具有以下性质: $\bar{A}(u)$ 越远离0.5,其值越小;反之越靠近0.5,其值越大。模糊集 \bar{A} 的模糊度与隶属度的关系如图1-1所示。

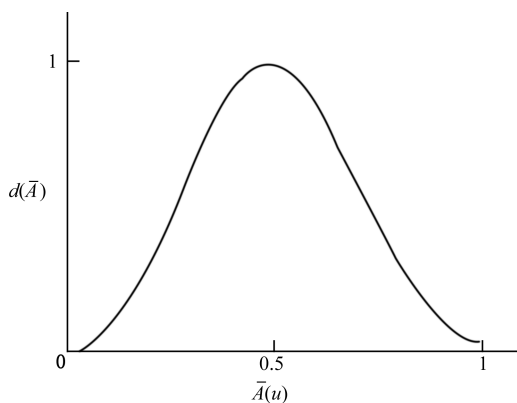


图 1-1 模糊集模糊度与隶属度的关系

2. 低消耗备件隶属函数的确定

“低消耗”是一个相对的概念,在实际工作中,无法精确地定义“低消耗”的标准,因此,可以采用一个模糊集合来定义低消耗备件。设论域 U 为所有的备件,论域 U 中的模糊集合 \bar{A} 代表低消耗备件, $\bar{A}(u)$ 为这个模糊集合的隶属函数,它代表某项备件隶属于低消耗备件的程度。模糊集合完全由它的隶属函数来表述,借助它才有可能对模糊集合进行量化,也才有可能利用精确的数学方法去分析和处理模糊信息。

本书采用模糊统计法确定低消耗备件的隶属函数。模糊统计法就是根据所提出的模糊概念进行调查统计。通过统计试验,得到每一元素隶属于模糊集合的程度,即求得模糊集合的隶属函数。求解步骤如下:

- (1) 给定论域 U 。
- (2) 就模糊概念向专家进行咨询,对咨询的结果进行统计,得到 U 的一个

运动着的、边界可变的普通集合 A^* 。在每次试验中, A^* 是一个确定的普通集合, 但在不同批次的试验中, A^* 的边界又可能是不同的, 因而把 A^* 作为 U 中一个可运动的普通集合。 A^* 与 \bar{A} 是有联系的, A^* 的每一次固定化, 都是对模糊概念做出的一个确定划分, 它表示概念的一个近似的外延。

(3) 做 n 次试验。对给定元素 $u_0 \in U$, 计算 u_0 对 \bar{A} 的隶属频率:

$$u_0 \text{ 对 } \bar{A} \text{ 的隶属频率} = \frac{u_0 \text{ 属于 } \bar{A} \text{ 的次数}}{\text{试验次数}} \quad (1-1)$$

随着试验次数的增加, 隶属频率会呈现稳定性。频率稳定所在的那个数称为 u_0 对 \bar{A} 的隶属度。

(4) 对所有 $u \in U$, 求 u 对 \bar{A} 的隶属度, 画出隶属度曲线, 根据曲线求出 \bar{A} 的隶属函数。

3. 低消耗备件界定

为将低消耗备件从众多备件中区别开来, 本书引用隶属度的概念。根据隶属度原则, 备件为“低消耗”的隶属度在 $[0,1]$ 区间内取值, 各单位可根据本单位的实际情况设定相应的阈值, 并依此制定“低消耗”的标准。

事物的发展往往具有一定的连续性, 历史消耗数据能反映出将来大致的消耗趋势。因此, 对低消耗备件的定义不能只针对某一年的消耗数据而论, 而是要将近几年的历史数据综合考虑。在对某项备件进行定义之前, 首先分别计算当年的“低消耗”隶属度 $\bar{A}(u_i) (i=1,2,\dots,n)$, 再对每年的隶属度赋予不同的权重 $\alpha_i (0 \leq \alpha_i \leq 1; i=1,2,\dots,n)$, 最后通过加权求和的形式来最终确定这项备件的“低消耗”隶属度, 即

$$Y = \alpha_1 \bar{A}(u_1) + \alpha_2 \bar{A}(u_2) + \dots + \alpha_n \bar{A}(u_n) \quad (1-2)$$

式中: u_i 为备件的年消耗量。

“低消耗”的隶属度取值范围为 $[0,1]$ 。当隶属度为 0 时, 其模糊度也为 0, 表示隶属度为 0 的备件可以明确地认为不属于“低消耗”的范围; 当隶属度为 1 时, 其模糊度为 0, 表示隶属度为 1 的备件可以明确地认为属于“低消耗”的范围; 当隶属度为 0.5 时, 其模糊度大致为 1, 表明此时“低消耗”的概念是最模糊的, 这也是低消耗备件概念界定的临界点。从这一临界点开始, 随着隶属度的增大, 模糊度随之减小, 表明隶属度在区间 $[0.5,1]$ 取值时, “低消耗”的概念越来越明确。因此, 根据模糊数学原理, 本书定义“低消耗”隶属度大于 0.5 的备件为低消耗备件。

4. 示例分析

选取某仓库的 10 项武器装备备件为研究对象，收集分析这些备件 2007—2010 年的消耗数据，如表 1-1 所示。按照本书提出的方法对这些备件进行分类。

表 1-1 2007—2010 年度备件消耗统计数据

年 份	年消耗备件数									
	备件 1	备件 2	备件 3	备件 4	备件 5	备件 6	备件 7	备件 8	备件 9	备件 10
2007	0	0	2	2	3	5	4	7	8	10
2008	0	0	1	4	5	4	7	7	10	12
2009	0	1	0	0	3	4	5	8	9	12
2010	0	0	1	2	4	7	6	9	9	14
年消耗均值	0	0.25	1	2	3.75	5	5.5	7.75	9	12

(1) 确定“低消耗备件”的论域。根据武器装备保障工作经验，年消耗量在 15 件或 15 件以上的备件通常可以确认不是低消耗备件，或者说，这类备件不是低消耗备件的判断是清晰的。这类备件对于“低消耗备件”这个模糊集合而言，它的隶属度为 0。本书以 $U = [0, 10]$ 作为“低消耗备件”的论域。

(2) 对 30 位长期从事武器装备保障工作的专家进行咨询，他们提出各自认为的低消耗备件的恰当范围，得到的统计数据如表 1-2 所示。

表 1-2 “低消耗备件”定义范围专家调查

年消耗数				
0~1.0	0~2.0	0~4.5	0~5.0	0~3.0
0~5.5	0~2.0	0~4.0	0~3.5	0~5.0
0~6.5	0~5.5	0~5.0	0~5.5	0~7.5
0~2.5	0~6.5	0~7.5	0~1.5	0~5.0
0~3.5	0~8.0	0~6.0	0~6.0	0~5.5
0~7.0	0~5.5	0~8.5	0~6.5	0~6.0

(3) 以步长 0.5，从论域 U 中选取 19 个点，并分别求出专家评判结果的隶属频率，如表 1-3 所示。

表 1-3 隶属频率

年消耗数	隶属次数	隶属频率	年消耗数	隶属次数	隶属频率
0	30	1	5.0	20	0.667
0.5	30	1	5.5	16	0.533
1.0	30	1	6.0	11	0.367

续表

年消耗数	隶属次数	隶属频率	年消耗数	隶属次数	隶属频率
1.5	29	0.967	6.5	8	0.267
2.0	28	0.933	7.0	5	0.167
2.5	26	0.867	7.5	4	0.133
3.0	25	0.833	8.0	2	0.067
3.5	24	0.800	8.5	1	0.033
4.0	22	0.733	9.0	0	0
4.5	21	0.700	—	—	—

表 1-3 中的隶属频率即隶属度。用 Matlab 软件描绘出隶属频率的曲线如图 1-2 所示。

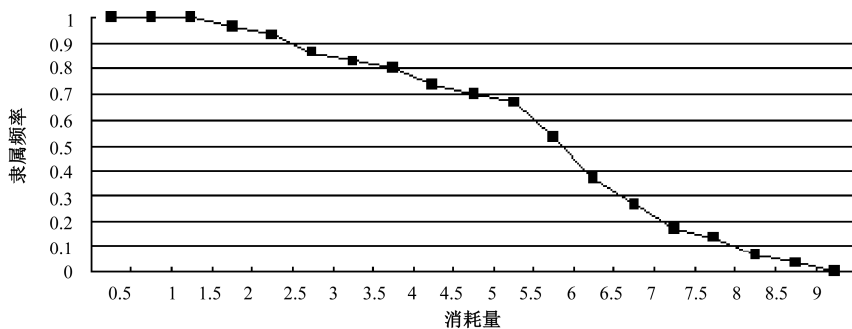


图 1-2 隶属频率曲线

对某些成套备件，由于年实际消耗量可能为小数，仅仅有消耗量为几个整数数据点的隶属度是不够的，因此，本书通过 Matlab 软件用多项式函数对隶属频率曲线进行拟合。拟合结果显示，当多项式函数的次数为 5 时，平均残差值为 0.032632，相关系数 $R = 0.9935$ ，拟合效果良好，如图 1-3 和图 1-4 所示。

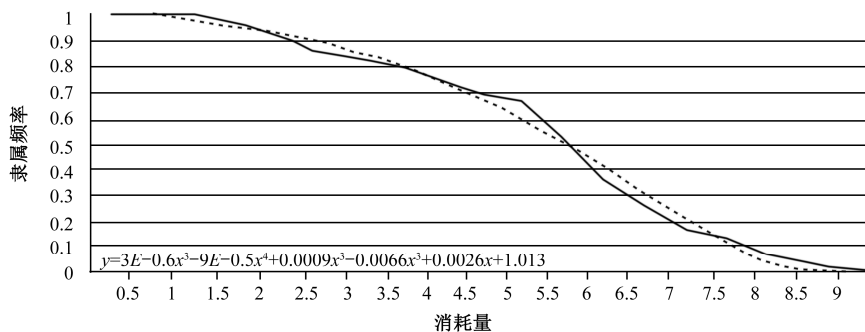


图 1-3 隶属频率拟合曲线

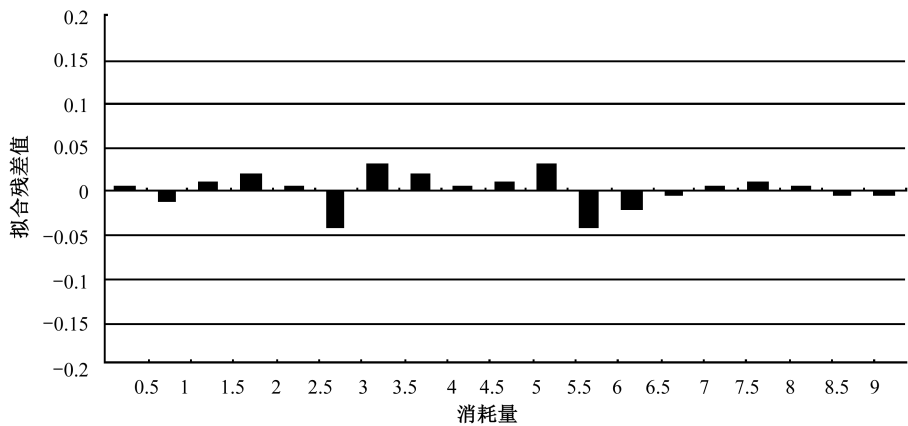


图 1-4 隶属频率拟合残差值

(4) 根据图 1-3 中的拟合曲线和式 (1-2)，计算各项备件的低消耗隶属度分别为 1,0.9899,0.9698,0.8794,0.7347,0.6271,0.4815,0.2224,0,0，然后根据计算结果对这些备件进行分类。计算结果显示：年消耗平均数低于 5.5 的备件共有 6 项；隶属度大于 0.5 的备件共有 6 项；隶属度大于 0.8 的备件共有 4 项；隶属度大于 0.9 的备件共有 3 项。

从以上结果可以看出：当阈值设定为 0.5 时，得到的低消耗备件为 6 项，与按照消耗平均数低于 5 的标准得出的结果基本一致；而当阈值设定为 0.9 时，低消耗备件共有 3 项，比阈值为 0.5 时减少了 3 项。这说明新的定义不仅覆盖了原定义的结论，同时，只要修改阈值，就可以重新设定低消耗备件的范围，这就便于做存储决策时，根据“低消耗”隶属度的不同，对低消耗备件划分不同的类别。

1.3 备件需求影响因素分析

武器装备在服役期间，主要是处于储存、准备、战备值班或维修等状态。通过对备件需求发生的因果关系进行分析，可知备件需求的产生主要来源于三个方面：①因故障发生而造成的备件更换，其占备件的大多数；②错误的拆卸造成的非故障零部件更换；③预防性维修备件的更换。影响武器装备备件需求的因素很多，武器装备备件需求影响因素可归纳为内在影响因素和外在影响因素两个方面，如图 1-5 所示。

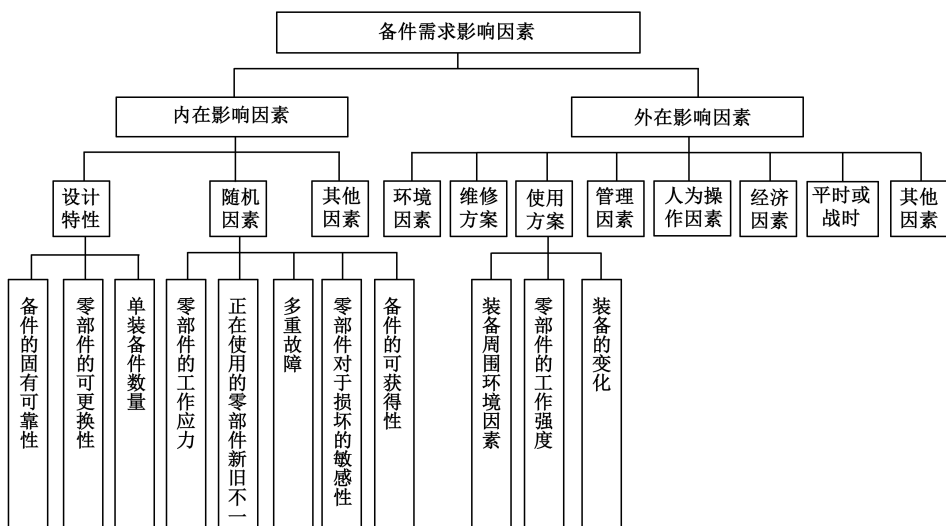


图 1-5 备件需求影响因素

1.3.1 备件需求内在影响因素分析

(1) 备件的固有可靠性。备件的固有可靠性，如基本故障率是备件在试验条件下得到的，它反映了备件本身的设计和制造水平，是备件的一种固有特性。其大小直接影响备件的需求，主要体现在非计划维修间隔时间或平均拆卸间隔时间上。

(2) 零部件的工作应力。同一构件在不同型号的武器装备上使用，或者虽在同一武器装备上但由于安装的位置不同，该构件受周围实际状况的影响也不同，发生故障的可能性及对备件的需求也不一样。例如，一种配电器，若用在某一装备上，会经常处于有害环境中，而用在另一装备上，其周围环境却比较清洁，则前者故障率必高于后者。

(3) 零部件对于损坏的敏感性。这主要是指在搬运、装配、维修及使用过程中，零部件由于非正常因素（人为差错或者技能水平与设计方面的问题）而受到损坏的可能性，而与该件的固有故障率无关。这种情况可以理解为“事故率”。如在装配中某电气元件被烙铁烧毁就是一个例子。

(4) 零部件的可更换性。这主要表现为备件拆装的难易程度，它与装备拆卸的程度有关，同时也与承修单位的维修水平和维修条件有关。影响可更换性的主要因素是产品的技术状态和互换性。

(5) 备件的可获得性。这主要表现为备件获得的难易程度，主要包括专用程

度、稀缺性等，有时用是否为标准件、专用件和延期交货时间来体现。

(6)正在使用的部件新旧不一。这是由于出现故障的部件将用新的备件替换，而未出现故障的部件将继续使用，新旧部件同时使用，相互间可能因为磨合或其他原因产生影响。

(7)多重故障。当一个部件发生故障后，由于条件变化，可能导致其他还可以使用的部件也发生故障，但其中的因素是多方面的，有一定随机性。

(8)单装备件数量。它表示一型装备设计方案中单台装备上相同备件的数量。如果某一种部件在武器装备中有若干个，并且是同时工作的，即相当于系统可靠性中的串联模型，那么这些同时工作的部件数量越多，备件的消耗量也越大。它与故障率的乘积表示部件的故障可能性大小。

(9)其他一些不确定、未可知的因素及偶发事件。

1.3.2 备件需求外在影响因素分析

(1)环境因素。武器装备所在地区的温度、湿度、盐度、盐雾、大气压力和霉菌的变化都会影响装备的使用可靠性，从而影响备件的需求。具体来说，温度对武器装备的影响分析如表 1-4~表 1-6 所示；湿度对武器装备的影响分析如表 1-7 所示；盐雾对武器装备的影响分析如表 1-8 所示；霉菌对武器装备的影响如表 1-9 所示；气压对武器装备的影响如表 1-10 所示。另外，环境因素不仅影响装备或部件性能，而且影响维修与操作人员的工作状态，从而加大人为损坏的可能，这在很大程度上增加了备件需求量。

表 1-4 高温对武器装备的影响及效应

影 响 类 型	后果或危害
材料热胀	装置活动部分产生黏合、卡死，紧固部分出现松动；接触不良；产品尺寸全部或部分改变；电子电路的稳定性改变
干裂	固体火药或药柱产生裂纹；有机材料开裂或出现裂纹
燃烧加快	爆炸物或推进剂的燃速增加
气体热胀	容器压力增大，封装弹药内压力增大
材料软化	结构强度减弱
油黏度降低	油液外流使连接处丧失润滑能力，轴承等传动部件损坏
设备过热	元件损坏，着火；焊缝融化、焊点开裂；炸药熔化和流化
化学分解和老化	材料性能改变，甚至丧失
金属氧化	接点接触电阻增大，表面电阻增加
其他	木制包装箱开裂；密封片变形、黏结；填充物和密封条易损坏

表 1-5 低温对武器装备的影响及效应

影 响 类 型	后果或危害
材料变脆	结构强度下降，并易断裂；橡胶、塑料硬化
材料收缩	龟裂、失去弹性，使减振或密封失效；配合间隙发生变化，使机械动作迟缓或停止；插头座、开关件等接触不良
油、润滑脂黏度增大， 防冻液冻结	轴承、开关等产生黏滞；机械动作的质量和精度下降，尤其是液压系统
水、汽冻结	设备和仪器动作停止，功能失灵；冷却水系统因堵塞而失效；光学仪器的观察性能下降
元器件性能改变	电阻、电容数值变化；电解电容器易损坏；石英晶体不振荡；蓄电池容量降低而使其寿命缩短
燃速下降	装药战术技术性能下降
其他	加速材料老化；加速金属的电化学腐蚀

表 1-6 变温对武器装备的影响及效应

影 响 类 型	后果或危害
材料剧烈膨胀与收缩	光学玻璃易碎裂；运动部件黏结或减速，零件变形或破裂；药柱产生裂纹；密封部件泄漏；涂层表面、材料界面开裂
凝水或结霜	电子或机械发生故障；仪器和仪表精度下降
其他	加速材料的老化；加速金属的电化学腐蚀

表 1-7 湿度对武器装备的影响及效应

影 响 类 型		后果或危害
高湿	水汽凝露	绝缘材料电阻和热性能降低，电气短路，降低电子元器件性能；玻璃或塑料光学元件的投射能力降低；密封设备材料的溶解和变化
	材料吸湿增强	材料膨胀、变形、物理性能降低，结构破坏；电性能变化；包装箱强度下降；炸药和推进剂变质
	加速金属氧化和腐蚀	润滑剂污染；活动零件易卡住；结构强度减弱；表面电阻增大，电接触不良
	加速化学或电化学反应	破坏有机涂层，材料化学和物理性能下降，加速金属腐蚀
	促进生物活动	促进霉菌生长，使材料产生或加剧受霉菌破坏
低湿	材料失水	木材、皮革和纺织物之类的材料因变干而发脆或粉化；易产生静电；加速材料干裂或龟裂；火工品性能变差
干湿交替	毛细管“呼吸”作用	加速材料反潮
	金属表面干湿交替	加速电化学腐蚀

表 1-8 盐雾对武器装备的影响及效应

影 响 类 型	后果或危害
化学反应	增大磨损，机械强度下降
锈蚀与腐蚀	电气性能变化，绝缘材料腐蚀
电解	产生电化腐蚀，结构强度减弱

表 1-9 霉菌对武器装备的影响及效应

影 响 类 型	后果或危害
分泌腐蚀液体	光学透镜表面薄膜侵蚀，金属腐蚀和氧化
吸附水分	导致其他形式的腐蚀，如电化腐蚀
霉菌	产生电化腐蚀，结构强度减弱

表 1-10 气压对武器装备的影响及效应

	影 响 类 型	后果或危害
高压	压缩	破坏结构；破坏密封；功能受影响
低压	膨胀	包装器材破裂甚至爆炸
	密封失效	电气性能变化，机械强度下降
	空气绝缘度下降	绝缘击穿，跳弧，出现电弧，产生电晕；漏气；出现臭氧
	散热不良	设备内部热量增加、聚集从而导致失效

环境因素通常分为自然环境因素和诱发环境因素，表现形式为组合环境因素。诱发环境因素又分为机械环境因素、核辐射、电磁辐射（包括静电放电）和大气污染。振动和冲击对武器装备的影响分析如表 1-11 和表 1-12 所示。

装备所遇到的环境因素大多是以组合形式出现的，如温度和湿度、温度和冲击或振动、湿度和霉菌等，这些成对环境因素之间的相互影响如表 1-13 所示。

表 1-11 振动对武器装备的影响及效应

影 响 类 型	后果或危害
机械影响	材料机械强度降低，磨损加剧，紧固件、连接件松动
电气影响	晶体管外引线、固体电路管脚折断；联结器、继电器、开关失效；电子插件性能下降；陀螺漂移增大甚至故障；仪表精度降低；导引头性能下降；粘层、键合点脱开，电路短路、断路

表 1-12 冲击对武器装备的影响及效应

影 响 类 型	后果或危害
机械影响	使产品结构因产生快速增长应力而失效，或者产生永久变形、折断和断裂；产品结构间的相对运动增大或减小构件间的摩擦或干扰而造成设备失效；加速材料的低周期疲劳
电气影响	改变绝缘强度，降低绝缘电阻，使磁场和静电场强度发生变化；电子器件引线、管脚、导线断裂，电路短路或断路

表 1-13 环境因素相互影响趋势

综合环境因素		综合影响趋势
高温	湿度	提高湿气浸透速度，加速湿气的锈蚀影响
	低气压	强化低气压因素的影响
	盐雾	增大盐雾所造成的锈蚀速度
	霉化	温度能促进霉化或微生物的生长，如果温度高于 71℃，则不能生长
	冲击、振动	强化冲击或振动的影响
低温	湿度	造成湿气汽凝，甚至出现霜冻或结冰
	低气压	加速因低气压造成的密封等的漏气
	盐雾	碱性盐雾的侵蚀速度
	冲击、振动	强化冲击或振动的影响
湿度	低气压	增大低气压的影响
	盐雾	高湿度可以冲淡盐雾浓度
	霉化	有助于霉化和微生物的生长
	臭氧	加大对塑料和弹性材料的侵蚀性作用

(2) 人为操作因素。人为差错的影响是指操作人员的行动在技术素质、心理素质，以及环境等因素的影响下产生的差错或误操作，人为的误操作也可能造成一些部件故障。

(3) 零部件的工作强度。零部件的工作强度包括构件的载荷情况和使用率等。比如同一部件安装在导弹上，超出正常使用频数和强度时，将影响某些零部件的故障率。超出正常范围的使用，一方面表现为使用连续时间过长或应力应变状况超出原设计条件，另一方面由于使用过少或没有使用，造成某些零部件变质或性能下降。

(4) 装备的变化。若武器装备的类型、数量发生变化，则其备件的需求量也相应变化。

(5) 经济因素。经济因素又称为设置成本，主要包括采购、运输、存储和维修备件所需要的各项费用。需要指出的是，考虑是否应该储备某种备件时，应该充分考虑该备件的单一成本和最小订货成本。显然，备件的设置成本越高，确定购置备件时越要慎重。

(6) 部件的维修方案。它规定了维修保障的工作内容，包括预防性维修与修复性维修项目、预防性维修类型及内容、维修间隔期、维修级别及修理策略等。其中，维修间隔期主要影响因预防性维修产生的备件需求；维修级别决定装备发生故障后在哪一级维修保障组织进行修理，将影响备件需求的产生和备件周转时间的长短；维修策略决定了备件发生故障后是进行维修还是报废，从而影响备件需求的产生。

(7) 备件保障中的管理因素。它主要是指因行政管理上的原因（如非装备保障的行政事务、指挥调度不当等）造成的备件供应延误，具体体现如平均管理延误时间。

(8) 平时或战时。这对备件需求的影响是显而易见的，战时表现为战场环境、作战对象、兵力对比等的影响，而平时则表现为训练环境和使用强度等的影响。

(9) 其他一些不确定因素及偶发事件。

1.3.3 影响因素的量化及规范化处理

在进行备件需求预测时，有时需要定量描述影响因素，以便用来进行分析计算，这就需要对影响因素进行量化处理。其主要是根据影响因素的种类和特点，深入现场或有关单位进行数据收集，然后选择不同的单位或量纲，对这些数据进行分析处理，最终可以得出影响因素的数值描述。容易发现，这些数据的量纲不尽相同，取值区间差异或大或小，影响因素标准可能天差地别。在进行备件需求预测时，分析和处理影响因素数据成为一个亟待解决的问题。采用类型归一化和无量纲化这两种方法可以很好地解决这个难题。

1. 类型归一化

备件需求预测建立的影响因素数据体系中，各影响因素的类型不同，通常可将其分为极大型（效益型）、极小型（成本型）、固定型（居中型）、区间型（值落在某区间 $[q_1, q_2]$ 内为最佳，越接近该区间越好）。预测前须将这些因素的类型归一化，以集结各属性值，否则无法用统一的标准判断它们对需求量的影响。

设所有影响因素集为

$$X = \bigcup_{k=1}^4 X_k \quad (k=1,2,3,4) \quad (1-3)$$

式中： X_k 分别为极大型、极小型、固定型和区间型因素集。

为了方便，将各类型因素集均转化为极大型因素集，则类型归一化的方法有：

(1) 对于极小型因素， $x \in X_2$ ，有

$$x' = M - x \quad (M \text{ 为因素 } X \text{ 的一个允许上界}) \quad (1-4)$$

$$x' = \frac{1}{x} \quad (x \neq 0) \quad (1-5)$$

(2) 对于固定型因素， $x \in X_3$ ，有

$$x' = \begin{cases} x - m & (m \leq x \leq a) \\ M - x & (a \leq x \leq M) \end{cases} \quad (1-6)$$

$$x' = \begin{cases} 1 & (x = a) \\ 1 - \frac{|x - a|}{\max\{x - m, M - x\}} & (x \neq a) \end{cases} \quad (1-7)$$

式中： M 、 m 分别为因素 x 的允许上界和下界； a 为固定值。

(3) 对于区间型因素， $x \in X_4$ ，有

$$x' = \begin{cases} 1 - (q_1 - x) / \max\{q_1 - m, M - q_2\} & (x < q_1) \\ 1 & (q_1 \leq x \leq q_2) \\ 1 - (x - q_2) / \max\{q_1 - m, M - q_2\} & (x > q_2) \end{cases} \quad (1-8)$$

式中： M 、 m 分别为因素 x 的允许上界和下界， $[q_1, q_2]$ 为因素 X 的最佳区间。

2. 无量纲化

各因素取值的量纲不同会给备件需求预测带来困难：一方面具有不同量纲的属性值无法做各种集结运算；另一方面即便量纲相同，如果各因素的取值区间差异很大，也会使某个因素所起的作用过大或过小，造成备件需求预测结果不合理。无量纲化是将因素实际值转化成因素指标值。从几何的角度可将无量纲化的方法归结为直线型无量纲化方法、折线型无量纲化方法、曲线型无量纲化方法三类。本书主要阐述一般的直线型无量纲化方法。

1) 阈值法

方法 1：

$$y_i = \frac{x_i}{\max_{1 \leq i \leq n} x_i}, \quad y_i \in \left[\frac{\min x_i}{\max x_i}, 1 \right] \quad (1-9)$$

该方法得到的指标值随着因素数值增大而增大,若因素数值均为正,则指标值不可能为零,因素最大值的指标值为 1。式(1-9)的坐标表示如图 1-6 所示。

方法 2:

$$y_i = \frac{\max x_i + \min x_i - x_i}{\max x_i}, \quad y_i \in \left[\frac{\min x_i}{\max x_i}, 1 \right] \quad (1-10)$$

该方法得到的指标值随着因素数值的增大而减小,适用于对逆因素数值进行无量纲化处理,即无量纲化和因素值转化同时进行。式(1-10)的坐标表示如图 1-7 所示。

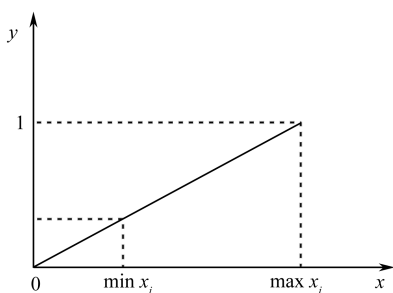


图 1-6 式(1-9)坐标表示

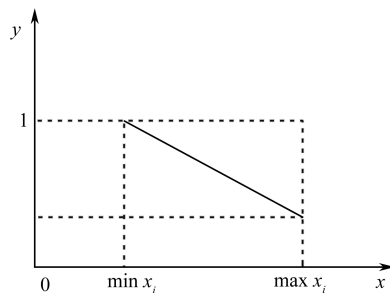


图 1-7 式(1-10)坐标表示

方法 3:

$$y_i = \frac{\max x_i - x_i}{\max x_i - \min x_i}, \quad y_i \in [0, 1] \quad (1-11)$$

该方法得到的指标值随着因素数值增大而减小,适用于对逆因素数值进行无量纲化处理,即无量纲化和因素值转化同时进行。式(1-11)的坐标表示如图 1-8 所示。

方法 4:

$$y_i = \frac{x_i - \min x_i}{\max x_i - \min x_i}, \quad y_i \in [0, 1] \quad (1-12)$$

该方法得到的指标值随着因素数值增大而增大,因素最小值的指标值为零,因素最大值的指标值为 1。式(1-12)的坐标表示如图 1-9 所示。

方法 5:

$$y_i = \frac{x_i - \min x_i}{\max x_i - \min x_i} k + q, \quad y_i \in [q, k + q] \quad (1-13)$$

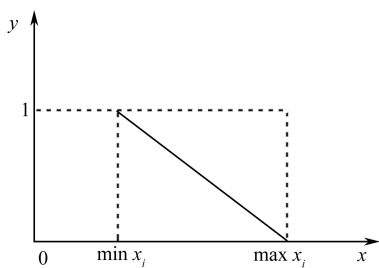


图 1-8 式 (1-11) 的坐标表示

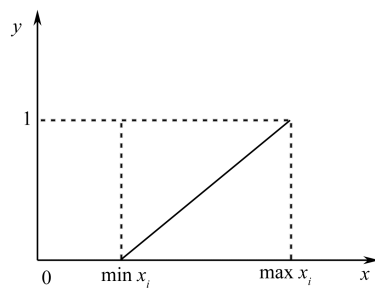


图 1-9 式 (1-12) 的坐标表示

该方法得到的指标值随着因素数值的增大而增大，因素最小值的指标值为 k ，因素最大值的指标值为 $k+q$ 。式 (1-13) 的坐标表示如图 1-10 所示。

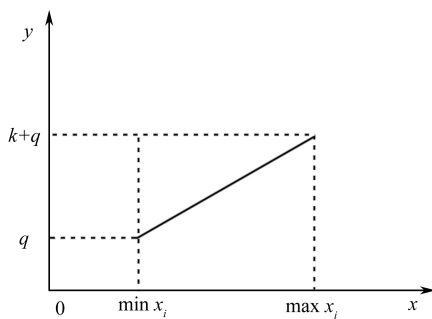


图 1-10 式 (1-13) 坐标表示

2) 标准化方法

$$y_i = \frac{x_i - \bar{x}}{s} \quad (1-14)$$

式中

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

3) 比重法

$$y_i = \frac{x_i}{\sum_{i=1}^n x_i} \quad \text{或} \quad y_i = \frac{x_i}{\sqrt{\sum_{i=1}^n x_i^2}} \quad (1-15)$$

1.3.4 基于粗糙集属性约简的主要影响因素确定

在进行备件需求预测时,输入矢量维数过多和训练样本量过大都会大大增加计算的复杂度,导致预测模型收敛速度慢,预测精度受限,从而降低模型的效率。而粗糙集(Rough Set, RS)属性约简可以从样本数据中挖掘信息,有效地从众多备件需求影响因素中优选出相关度较高的影响因素,从而使备件需求影响因素的选取更加全面、科学和合理,有效地提高预测速度和精度。

1. 粗糙集属性约简

粗糙集理论作为一种处理不精确、不一致、不完整等各种不完备信息的有效工具,一方面得益于它的数学基础成熟、不需要先验知识,另一方面在于它的易用性。由于粗糙集理论创建的目的与研究的出发点是直接对数据进行分析 and 推理,从中发现隐含的知识,揭示潜在的规律,因此它是一种天然的数据挖掘或知识发现方法。

在粗糙集理论中,知识基于对对象分类的能力。对任一信息系统 $K = (U, R)$ (其中, U 为所研究对象的全体,称为论域, $U = \{x_1, x_2, \dots, x_n\}$; R 为属性集, $R = \{r_1, r_2, \dots, r_n\}$), 任一 $P \subseteq R, P \neq \emptyset$, 则全集 U 上的元素关于 P 的等价类构成一种不可分辨关系, 记为 $\text{ind}(P)$ 。以 $f(x, a)$ 表示元素 x 关于属性 a 的取值, 若 x_{i1} 和 x_{i2} 关于 P 不可分辨, 则对任一 $r_i \in R$, 满足 $f(x_{i1}, r_i) = f(x_{i2}, r_i)$ 。

若存在不可分辨关系 P , 则由 P 形成的等价类可表示为 $U / \text{ind}(P) = \{X_1, X_2, \dots, X_n\}$ 。对于集合 $X \subseteq U$, 若 X 能表示成这些等价类的并集, 则称集合 X 相对于关系 P 可精确定义, 否则, 只能通过下近似和上近似来表示。集合 X 关于 P 的下近似和上近似分别用 $P_-(X)$ 、 $P^+(X)$ 来表示, 满足: $P_-(X) = \bigcup \{X_i | X_i \subseteq X\}$ 和 $P^+(X) = \bigcup \{X_i | X_i \cap X \neq \emptyset\}$ 。下近似 $P_-(X)$ 又称为 X 的 P 正域, 记为 $\text{pos}_P(X)$ 。

设 P 和 S 为 U 中的两个等价关系, $U / S = \{X_1, X_2, \dots, X_n\}$, S 的 P 正域记为 $\text{pos}_P(S)$, 即 $\text{pos}_P(S) = \bigcup_{i=1}^n P_-(X_i)$ 。若存在 $r \in R$ 使 $\text{pos}_P(S) = \text{pos}_{P-\{r\}}(S)$, 则认为 r 为 P 中 S 是可省略的, 即 $P - \{r\}$ 为 P 的 S 相对简化。

2. 示例分析

某型机载武器装备主要处于储存保管、战备值班、挂机飞行、维修等状态,

其备件需求量受自身可靠性、通电检测时间、战备值班时间等多种因素影响。试从中分析确定影响备件需求的主要因素。

经过对该型武器装备使用与保障情况的分析,影响该型武器装备备件需求的因素有固有故障率(导弹定检故障率)、通电检测次数、拆卸次数、战备值班时间、挂机飞行时间、恶劣工作天气(室外温度高于 30°C 或低于 -5°C 、相对湿度高于 70%等)占当年工作日的比例、业务素质好的保障人员所占的比例、年度重大任务次数,组成属性集合 $R = \{A, B, C, D, E, F, G, H\}$ 。对每个属性设定阈值,1 表示不小于阈值,0 表示没有达到标准。表 1-14 列出了备件需求影响因素信息。由此可知: H 的属性值在 10 年跨度里只有一个不同,这说明每年的重大任务变化不大,对备件的影响程度基本稳定,对预测不会有太大影响,可以把它去掉; D 的属性值和 E 的属性值基本相同,所以只保留一个,假设保留 D ,删除 E 。

2000—2009 年备件需求影响因素组成集合 $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ 。

表 1-14 备件需求影响因素信息

年 份	影响因素量化值							
	A	B	C	D	E	F	G	H
2001	0	1	0	0	0	1	0	0
2002	0	0	1	1	1	1	1	0
2003	1	1	0	0	0	0	0	0
2004	0	0	0	0	0	1	0	0
2005	1	1	1	1	1	1	1	1
2006	0	1	1	1	1	0	0	0
2007	1	1	1	0	1	1	1	0
2008	0	1	1	1	1	1	1	0
2009	0	0	0	0	0	0	0	0
2010	1	1	1	1	1	1	1	0

下面对表 1-14 中的数据进行约简。其中 $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $R' = \{A, B, C, D, F, G\}$, 则 $\text{ind}(R')$ 关系有以下等价类:

$$U / \text{ind}(R') = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5, 10\}, \{6\}, \{7\}, \{8\}, \{9\}\}$$

$$U / \text{ind}(R' - A) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5, 8, 10\}, \{6\}, \{7\}, \{9\}\} \neq U / \text{ind}(R')$$

$$U / \text{ind}(R' - B) = \{\{1, 4\}, \{2, 8\}, \{3\}, \{5, 10\}, \{6\}, \{7\}, \{9\}\} \neq U / \text{ind}(R')$$

$$U / \text{ind}(R' - C) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5, 10\}, \{6\}, \{7\}, \{8\}, \{9\}\} = U / \text{ind}(R')$$

$$U / \text{ind}(R' - D) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5, 7, 10\}, \{6\}, \{8\}, \{9\}\} \neq U / \text{ind}(R')$$

$$U/\text{ind}(R' - F) = \{\{1\}, \{2\}, \{3\}, \{4, 9\}, \{5, 10\}, \{6\}, \{7\}, \{8\}\} \neq U/\text{ind}(R')$$

$$U/\text{ind}(R' - G) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5, 10\}, \{6\}, \{7\}, \{8\}, \{9\}\} = U/\text{ind}(R')$$

所以属性 C 、 G 可省略，而属性 A 、 B 、 D 、 F 不可省略，这样就把原来 8 个影响因素精简为 4 个。

1.4 备件需求层次性分析

备件需求主要是由武器装备维修保障活动引起的。按照备件供应保障过程来分，维修保障任务引起备件需求，并通过备件申请、筹措、储备及供应形成具有闭环特征的保障过程反馈关系。从备件筹措、订购与储备的方式来看，目前订购备件时是按照保障组织级别、某型装备、装备子系统、零部件等依次对备件品种和数量进行采购的；根据备件订购的层次性不同，将其划分为单个备件（或更换单元）、整个装备的和保障组织的备件订购标准。对应地，可按备件需求层次的不同，将其分为单个备件、单装备系统及维修保障组织 3 个层次来研究备件需求，如图 1-11 所示。

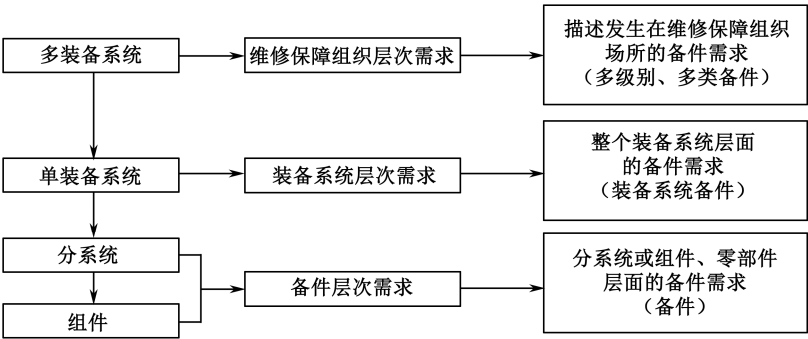


图 1-11 备件需求层次结构

1.4.1 备件层次需求分析

备件层次需求是指装备中分系统、组件或部件的备件需求，对应现场可更换单元（Liner Replaceable Unit, LRU）、车间可更换单元（Shop Replaceable Unit, SRU）、报废单元（Discardable Unit, DU），它们占据装备系统层次需求的大部

分。例如,对于海军军械装备保障,基层级(军械技术保障大队或修理所)以整装换件修理为主,对应的就是系统级备件需求,LRU 层次;军械修理厂以整装中修换件修理为主,对应的为部分 LRU 和 SRU;生产厂以换件修理与原件修复性修理相结合,对应的主要是 SRU。

1.4.2 装备系统层次需求分析

装备系统层次需求是指将装备系统看作一个整体进行描述,分析系统维修和系统备件需求,即按装备型号的划分确定备件需求。装备系统层次备件需求以满足系统自身需求来实现,这也是采用系统分析法进行备件优化配置的基础。以更换装备中的 LRU 为例,当满足备件需求时系统得到恢复,系统战备完好性得以保持,否则直接影响装备的战备完好率。在装备维修过程中,更换下来的故障备件会产生维修需求并按照自下而上的顺序从基层级维修保障组织到生产厂级维修保障组织传递,直到故障备件被鉴定为废弃或修复为止。在故障备件传递的过程中,各保障组织可能产生对应的备件需求,或者因故障备件被修复而不引起备件需求,具体情况根据各保障组织维修能力而定。

1.4.3 维修保障组织需求分析

维修保障组织是实施维修保障活动的主体,它包括装备部署、维修及备件存储等多个单位,其基本要素为阵地、修理间和库房,因而,保障组织可描述为一系列站点的集合。确定装备在维修保障组织中的备件需求,首先需要对备件需求在各维修级别传递的过程进行描述。目前,武器装备维修体制通常情况下属于三级维修,即基层级、中继级和基地级维修。有些武器装备实行的是基层级和基地级两级维修体制。

1.5 备件需求特点分析

1.5.1 备件需求特点

武器装备备件需求是指为了维护国家安全利益,军队在一定价格条件下购买的,用于实施和保障武器装备、武器系统及其配套的军事技术装备与器材的零部

件的品种和数量,是为了实现一定国防目的而产生的。其具有以下特点:

(1) 品种繁多,科技含量激增。目前,我军武器装备型号多、时间跨度长,设计与制造水平差异很大,备件标准化、通用化、模块化水平还比较低,特别是部分老型号装备,可靠性、维修性与保障性相对较差,自身的保障设备众多,备件需求呈现数量大、品种多的特点。此外,有些武器装备结构复杂,自动化程度高,是多项高新技术综合运用的复合体,这在很大程度上决定了装备所需备件的品种型号繁多。部件类别由原来以机械部件为主逐渐转变到以电子部件为主,机械、电子、光学、光化学部件齐全的格局,也使得备件科技含量激增,保养、检测难度加大。

(2) 影响因素多,波动性大。任何商品的需求由于其影响因素的变动都会存在一定程度的波动。很多武器装备备件由于使用目的的特殊性,影响需求量的因素不确定,因而波动性显得更加明显。其主要表现在需求的规模波动和结构波动两个方面。规模波动也就是需求总量的波动,会呈现收缩和扩张两种趋势;而结构波动则较为复杂,主要呈现在备件品种不断更换方面。

(3) 需求规律性差,难以准确预测。平时备件消耗主要源于预防性维修和修复性维修,装备的故障原因、机理和模式基本是可以预见的,有一定的规律性;而战时备件消耗主要源于装备的战斗损伤和其他一些不可预见的因素,其不确定性、随机性很大,难以事先预计。即使在相同的作战时间、类型及战场条件下,同类备件的消耗也可能不同。目前我军装备呈现出非均衡发展的势头,一些特种武器装备作用大、数量少、失效机理不明确,其备件需求量不大;但其价格高,难以采购,不可或缺,呈现“低密度”的特点,无规律性可言,其备件需求预测就更加困难。

1.5.2 常消耗备件需求特点

在本书中,常消耗备件主要指武器装备上装配数量较多、工作频率较高或寿命较短的元器件、零件、组件、部件或分系统,其最大特点是备件需求量相对较大,而且与装备的固有特性和使用因素有着直接的关系。

在武器装备的服役早期,装备各个(分)系统或部件处于磨合期,且要适应新的工作环境和储存条件,所以故障率较高,但随着时间的增加而迅速下降,故障主要是由部件设计、加工工艺、材料选择不合理或操作人员误操作造成的。因此,这一时期需要配备一定数量的备件。

在武器装备的服役中期,装备从磨合期进入偶然故障期,装备故障率较低且

基本恒定,装备故障主要是由偶然因素引起的,此期间发生的故障是随机的,备件消耗表现为自然消耗。在装备形成战斗力后,部队的训练常态化,演习和战备任务加重,还有一些环境和人为因素影响,这都会增加备件需求。因此,这一时期的备件需求往往具有一定的波动性,但总的趋势相对稳定。

在武器装备服役后期,由于装备(分)系统或零部件老化、衰变、退化、材料蠕变、疲劳等原因,装备故障率随时间的延长而逐渐增加,因此,这一时期的备件需要量逐渐增多,直至装备报废。

常消耗备件的需求与装备固有可靠性、装备使用强度、工作环境等因素之间存在密切的关系,在装备数量变化不大的情况下,通常存在一定的规律性。而且随着武器装备使用时间的延续,在服役中期后续阶段已经积累了大量的备件需求数据,采用基于统计数据的预测方法进行备件需求预测是可行的。

1.5.3 低消耗备件需求特点

随着现代科学技术的迅猛发展,武器装备的质量不断提升,再加上一些设计和使用方面的因素,导致某些武器装备备件的消耗量相对较少,以及消耗规律不容易确定,这类备件称为低消耗备件或不常用备件、低需求备件。该类备件多为武器装备的核心部件或关键件,一般具有专用性强、功能关键、可用性要求高、价值昂贵、使用频率低、订购周期长等特点。

对大部分时间处于储存状态的武器装备,在对其某些低消耗备件历史需求数据统计时发现,其备件需求时间序列具有间断性特点,即时间序列中非零需求是稀疏的,非零需求之间有很多零需求,而这些非负整数的出现表现出很强的随机性。这就是通常所说的间断需求件,也称为慢速流动部件,它们一般不发生故障,在历史需求数据中往往表现出一定数量的零值。

武器装备发生故障时,如果不能及时提供所需备件,将严重影响装备的战备完好性,而过度库存又将占用大量的费用。这种间断性的低消耗特征使得预测非常困难,用一般的连续性预测方法进行预测,精度很低。目前,针对这类备件需求预测方面的研究还相对较少。间断需求件需求预测的难点在于需要同时预测需求发生的概率和需求量。

第 2 章

武器装备备件需求预测与 配置优化基础



预测是以准确的调查统计资料和数据为依据，从研究对象的历史、现状和规律出发，运用科学的方法，对研究对象的未来发展前景的测定。经过近几个世纪的发展，科学技术极大地提高了人类的认识能力，也同样促进了人类对充满不确定性的未来的预测能力。对于自然系统，人类已经可以在一定的时间尺度下相当精确地预测它的发展变化。然而，在武器装备备件保障领域，预测仍然是一个比较困难的问题。因此，针对不同的备件保障情况及要求，可选择不同性质及类别的方法进行预测。

2.1 备件消耗模式分析

备件消耗模式是指备件消耗随时间变化的规律。归结起来通常有备件消耗的线性趋势模式（见图 2-1）、备件消耗的非线性趋势模式（见图 2-2）、备件周期性的消耗变化模式（见图 2-3）、备件线性趋势与周期性变化的组合模式（见图 2-4）四种。

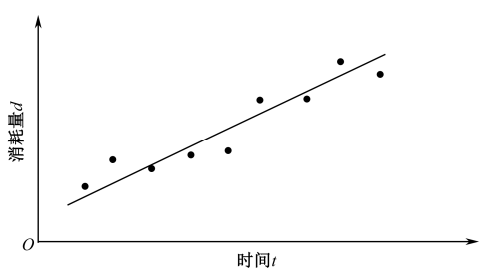


图 2-1 备件消耗的线性趋势模式

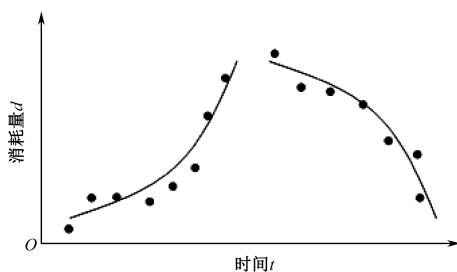


图 2-2 备件消耗的非线性趋势模式

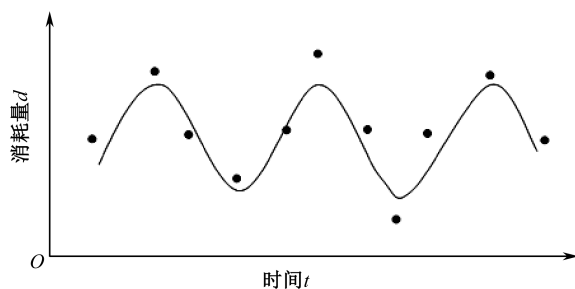
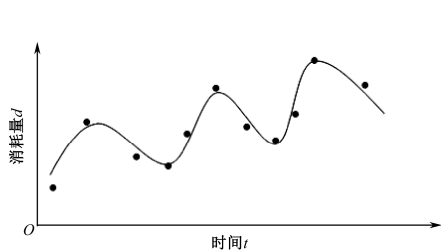
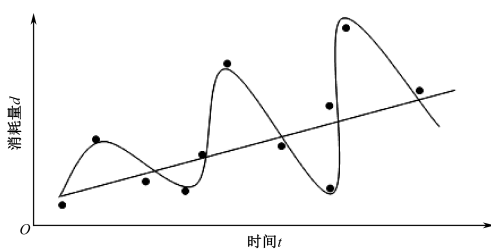


图 2-3 备件周期性的消耗变化模式



(a) 相加组合模拟



(b) 相乘组合模式

图 2-4 备件线性趋势与周期性变化的组合模式

2.2 备件需求预测的基本原则

从表面上看，事物的变化和发展似乎是杂乱无章的，是由各种偶然因素决定的。如果能够从已发生的事实中认识一种事物发展变化的规律性，就可以利用这

一规律性对事物的发展前景进行预测，并可期望达到和实际情况相符合的结果。预测是在掌握相关信息的基础上，运用数学、统计学、计算机、工程技术及经验分析等定性定量的方法，研究事物未来发展及其运行规律，并对其各要素的变化趋势做出估计描述与分析的一门学科。科学地预言尚未发生的事物是预测的根本目的和任务，无论对于个体或对于组织，在其制定规划策略等面向未来的决策过程中，预测是必不可少的重要环节，是科学决策的重要前提。

在进行武器装备备件需求预测时，应该遵循以下基本原则。

(1) 延续原则。事物的发展都具有一定的延续性，备件需求的变化发展也不例外。未来的备件需求是从今天的现状发展而来的，是今天的延续和发展。根据延续原则，过去和现在的备件需求中存在的某种规律，在将来的一段时期内将继续存在。延续性的内容包括时间和结构的延续性，其为根据样本资料进行外推提供了理论依据。武器装备发展普遍具有同源性和继承性，其（分）系统或零部件延续前一型号的性能要求和结构功能，因此备件品种具有很强的延续性；而部队训练计划和任务在一定时期内不变，可依据备件需求规律外推下一时期备件的需求量，因此备件需求变化也具有延续性。

(2) 类推原则。类推原则是指利用预测对象与其他事物的发展变化在时间上前后不同，在表现形式上有相似之处的特点，把发展事物的表现过程类推到预测对象上，对预测对象的前景进行预测。对武器装备备件需求预测来说，不同的备件需求所遵循的发展规律有时是相似的，利用这种相似性，可以根据已知的某一备件需求的发展规律类推出另一未知的备件的需求。比如在环境、使用条件、使用场合和管理水平等因素相近的前提下，相似备件的需求也是相近的。

(3) 相关原则。在进行武器装备备件需求预测时，会发现某些影响因素与备件需求量之间往往存在一定的相互联系和相互影响，即影响因素与某些备件需求量之间存在一定的相关性。相关性有多种表现形式，其中最重要的是因果关系。因果关系的特点是原因在前，结果在后，并且原因和结果之间常常有类似函数关系的密切联系。通过对一组影响因素与备件需求量的分析研究，确定出原因和结果后，就可利用这些实际统计资料来建立数学模型，从而进行备件需求预测。

(4) 统计规律性原则。在影响武器装备备件需求的众多因素中，有的还没有认知或者很难量化，很难准确描述其函数公式，更无法将其纳入数学模型，从而很难了解它对备件需求的影响和作用。这些影响因素的作用之和具有不确定性，使得备件需求呈现随机变化的形式，不可能找出备件需求的确定性变化规律，只能利用观测数据得出其统计规律，以一定的概率对备件需求的未来变化做出预测。

(5) 反馈原则。一般来说，备件需求的预测值不可能正好等于实际观测值，

二者之差就是预测误差。误差是很重要的信息，其大小和符号说明了数学模型和客观实际相结合的程度。为此，在进行武器装备备件需求预测时应当利用预测误差对预测模型或参数进行修正，尽量使模型符合实际情况，预测出比较准确的消耗量，从而积极推进精确化保障。反馈是评价和检验预测效果的有效手段，也是进行预测方法优化和改进的必要步骤，能产生很可观的军事效益和经济效益。

2.3 备件需求预测的一般程序

任何预测都要按照一定的步骤和程序进行，武器装备备件需求预测也不例外。为明确预测思路，提高预测效率，结合预测对象的特点，归纳总结了武器装备备件需求预测程序，如图 2-5 所示。

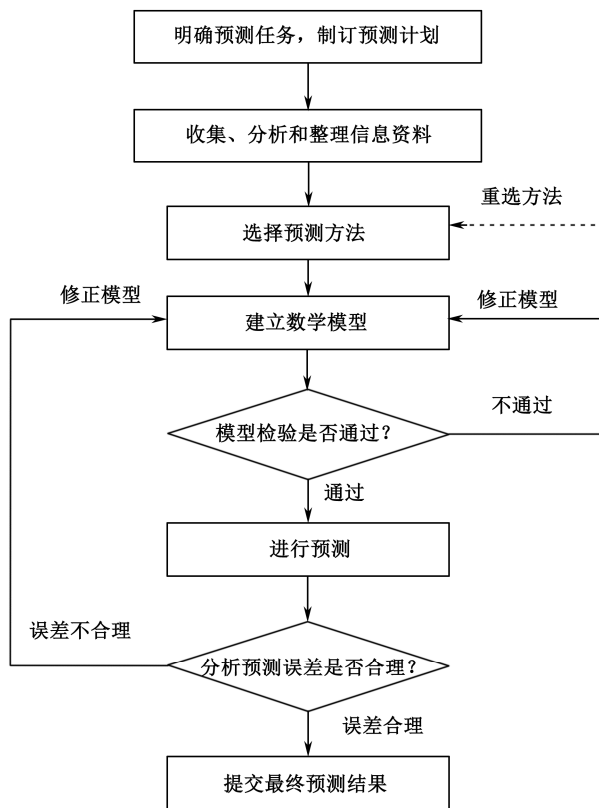


图 2-5 备件需求预测程序

(1) 明确预测任务, 制订预测计划。确定预测任务是从决策与管理的需要出发, 紧密联系实际需要和可能, 确定武器装备备件需求预测要解决的问题。预测计划是根据预测任务制订的预测方案, 包括预测的内容、项目, 所需要的资料, 准备选用的预测方法, 确定预测精度等。

(2) 收集、分析和整理信息资料。进行预测需要有大量的历史数据, 掌握与预测目的、内容有关的各种历史资料, 以及影响未来发展的现实资料, 即要从多方面收集资料。系统的、准确的相关信息及其他有关的原始资料和数据, 是开展预测分析的前提条件。资料按来源不同有内部资料和外部资料之分。内部资料是反映该单位与预测项目相关的统计资料、分析研究资料。外部资料是从单位外部收集到的统计资料和信息。依照直接有关性、可靠性和最新性的标准来筛选与预测项目有密切关系的资料。收集的信息和资料, 既要完整、全面, 又要重视资料的可靠性。有些信息还需进行加工、整理、归纳、鉴别、“去伪存真, 去粗存精”, 找出各因素之间相互依存、相互制约的关系, 并从中发现事物发展的规律, 作为预测的依据。武器装备备件消耗数据具有特殊性, 不但部分数据不够齐全, 而且大多具有保密性。通过调研收集资料存在很大的难度, 数据的准确性和可靠性也不太高, 这时对资料的整理与分析就显得格外重要。

(3) 选择预测方法, 建立数学模型。在占有资料的基础上, 对历史数据进行识别, 研究数据的模式, 进一步选择适当的预测方法并建立数学模型, 这是预测准确与否的关键步骤。对定性或定量预测方法的选择, 应根据掌握资料的情况而定。当掌握的资料不够完备、准确度较低时, 可采用定性预测方法, 建立设想的逻辑思维模型, 并拟订预测的调查提纲。如对新型武器装备备件进行预测时, 由于缺乏历史统计数据, 可凭预测者的经验进行预测。当掌握的资料比较齐全、准确度较高时, 可采用定量预测方法, 运用一定的数学模型定量分析研究。为充分考虑定性因素的影响, 在定量预测基础上要进行定性分析, 经过调整才能定案。如何针对武器装备备件需求预测的具体情况选择合适的数学模型, 是关系到预测准确程度的一个关键问题。

(4) 模型检验, 进行预测。模型建立后必须经过检验才能用于预测。模型检验主要包括考察参数值在理论上是否有意义, 统计显著性如何, 模型是否有良好的超样本特性。不同类型的模型检验的方法和标准不同。一般评价模型优良的基本原则是理论上合理、统计可靠性高、预测能力强、简单适用。若检验不通过, 则需要修正模型或者重新选择方法; 对于通过检验的模型, 按一定的要求对自变量赋值, 可算出因变量对应的预计值。

(5) 分析预测误差, 评价预测结果。分析预测值偏离实际值的程度及其产生

的原因。如果误差未超出允许的范围,即认为模型的预测功效合乎要求;否则,就需要查找原因,对模型进行修正和调整。

(6) 修正并提交预测结果。对于原用定量方法进行的预测,常常由于某些因素的数据不充分或无法定量而影响预测的精度,这就需要用定性方法考虑这些因素,并修正定量预测结果。至于原用定性方法预测的结果,往往也需应用定量方法加以修正、补充,使预测结果更接近实际。针对武器装备备件需求预测需要考虑下一年或几年内装备数量、训练计划、作战任务等是否有明显变化的情况,可通过引进不确定性系数来优化预测结果,即用定量计算和定性分析相结合的方式得到预测结果并最终提交。

2.4 备件需求预测的方法

20 世纪下半叶是预测发展迅速的一个时期,预测理论经历了几个重要的发展阶段。科学在近几个世纪的发展,极大地提高了人类的认识能力,也同样促进了人类对充满不确定性未来的预测能力。针对不同的环境及要求,预测理论的研究者提出了大量的方法来解决这一难题。

2.4.1 定性备件需求预测方法

定性预测主要是研究和探讨预测对象在未来所表现的性质,如事物发展的总体趋势,事件发生和发展的各种可能性及其造成的影响等。定性预测方法又称为主观预测方法,是依靠熟悉业务知识、具有丰富经验和综合分析能力的人员与专家,根据已掌握的历史资料和直观材料,考虑各种影响因素,对未来发展趋势与变化做出性质和程度上的判断,然后通过一定的形式综合各方面的意见,对事物未来表现的性质进行推测和判断。这类方法主要用于对预测对象的历史数据掌握不多或影响因素复杂的情形。它不采用复杂的数学公式,简便易行。其缺点是预测的准确程度依赖于人的经验和主观判断,缺乏数量上的准确描述。

定性预测方法目前仍得到广泛应用,它主要是在没有或缺少历史统计数据的情况下,依据人们的主观经验进行预测。

(1) 主观概率法。客观概率是根据事物发展的客观性统计出来的一种概率。主观概率不同于客观概率,它是预测者对未来事件发生可能性大小的一个经验信念度量。主观概率法是先由预测专家对预测事件发生的概率做出主观的估计,然

后计算它们的平均值，并以此作为对事件预测的结论。

概率统计法用于备件需求确定和优化已经有近百年的历史(最早是由管理学家 Harris 在 1918 年推导出了库存模型)。它主要是根据装备的故障规律，通过概率论与数理统计理论、排队论等知识建立备件的消耗模型、各级备件库存模型等。

(2) 抽样调查推算法。抽样调查推算法是根据随机原则，在所要调查的总体范围内，抽取一部分具有代表性的样本，通过对样本的调查来推算总体的一种方法。

(3) 名义团体技术又叫名义群体法 (Nominal Group Technique, NGT)，即逐步形成群体意见的方法，它是对一群体决策问题，发动群体成员提意见，然后按一定程序集结成员的意见，以做出群体的判断。这种方法适合较小的群体，以 5~9 个成员为宜。整个过程通常需要 60~90min。NGT 也具有统计的群反映特征，预测精度取决于对专家经验合理地收集和归纳。一般来讲，专家经验系统较难完整地进行总结归纳。因而，这里专家推理是辅助方法。

(4) Delphi 法。Delphi 法是决策、预测和技术咨询的一种有效的方法。它与 NGT 有相似之处，都是对复杂的决策问题通过征求和集结群中成员的意见去做出群的判断。两者的区别在于：一是 Delphi 法群中成员的人数一般比 NGT 多，以 20~50 人为宜，并且不要求成员面对面地接触，仅靠成员的书面反映；二是 NGT 通常可在 1~2h 内完成，Delphi 法一般需要数周甚至更长的时间。Delphi 法有三个主要特征，即匿名反映、迭代和受控的反馈、统计的群反映。

(5) 综合判断法。这类方法是由熟悉业务又能预见发展趋势的管理人员，凭借个人的知识和经验对未来的发展趋势做出预计和推测。对他们的预计和推测，一般采用简单的算术平均或加权平均的方法加以综合，从而计算出其预测值。

定性预测方法还包括典型调查推算法、类推法和相关因素分析法等多种方法。一般来说，定性预测方法的预测精度较低，但是，随着科学技术的发展，定性预测方法也在不断完善，不断提高其科学性和准确性。当对预测对象所掌握的历史性统计资料较少，又缺乏经验或由于其他原因无法应用定量预测方法时，往往采用定性预测方法来进行预测。同时，定性预测方法还经常作为定量预测方法的补充来进一步完善预测结果。

2.4.2 定量备件需求预测方法

定量预测方法是指对预测对象未来的数量表现加以确定的方法。它主要依靠历史统计资料，运用数学方法建立可以表现变量之间数量关系的模型，并利用这一模型来预测对象在未来可能表现的数量。其优点是注重事物发展在数量方面的

分析,注重发展变化的程度上数量的描述,预测主要依据历史统计资料,较少受主观变化因素的影响,可采用计算机辅助处理预测数据的计算。其缺点是对信息资料的质量和数量要求较高。有时结合主观判断,可使定量预测更经济合理。

(1) 因果回归预测方法。该方法是从事物变化的因果关系出发,寻求它们之间依存关系的一类预测方法。它的主要特点在于有清晰的模型结构。正确的模型可以很好地理解变量间的相互关系。这类方法于20世纪60年代在预测领域占主导地位。但是由于其所需的历史数据较多,而且在实际情况中,一个事物变化的全部原因是不可能完全得到的,因此影响了这类方法的预测能力。尽管如此,因果关系的思想却始终深刻地影响着预测理论的发展。

(2) 时间序列预测方法。时间序列是指把历史统计资料按时间顺序排列起来得到的一组数字序列。时间序列预测方法是用反映被测事物过去和现在变化规律的观测数据构造时间序列模型,然后借助模型进行外推以预测未来。它主要用于分析影响事物的主要因素比较困难或相关变量资料难以得到的场合,实际上是把影响变量的各种因素全部归结为时间变量。

(3) 仿真法。实际情况中,零部件的寿命分布并不一定明确,使备件保障进行决策较为困难。利用仿真法可以解决这方面的问题,其中蒙特卡罗方法是一种能够求解非确定分布备件最优储备量的随机模拟算法,运用该算法直接模拟部件和备件实际使用过程,通过对模拟结果进行统计分析获得备件最优储备量。

(4) 组合预测方法。组合预测方法是将不同的预测方法进行适当的组合,综合利用各种方法所提供的信息,以适当的方式得出组合预测模型和最佳预测结果,从而尽可能地提高预测精度。只要组合适当,这一目的完全可以实现。显然,组合预测方法比单一的预测方法更为科学,因为对同一预测问题,不同的预测方法提供不同的有用信息,如果简单地将预测误差较大的一些方法舍弃,将会丢失一些有用的信息,影响预测的精度。

组合预测方法主要有以下优势:

① 能从不同的角度、不同的样本数据、不同的模型得到系统不同的信息,能够博采众长,更全面地把握系统,达到提高预测精度与增加预测稳定性和结果可靠性的目的。

② 在典型的单项预测中,对于预测的开始阶段,选择正确的预测模型或方法很困难,一般采用所有模型来尝试、比较,但很难找到“最佳”模型。组合预测方法使得模型选择问题相对容易一些,许多经验(包括许多长期预测)表明,通过组合不同的模型,预测精度通常比单个模型有所提高,而不用去找“真实”或“最佳”模型,而且组合模型对于数据结构的变化具有更强的鲁棒性。

③ 可弥补单项预测方法不准确的缺陷,减少风险性。事实上,组合预测方法本质上是将各种单项预测看作代表不同信息的片段,通过信息的集成,分散单个预测特有的不确定性和减少总体不确定性,从而提高预测精度。组合预测方法相对于单项预测方法具有更高的预测性能,出现预测误差的风险更小。

组合预测方法可以规避预测风险,是提高预测精度的最佳方法之一。它适用于武器装备备件消耗具有时间波动性和随机波动性的情况。在组合预测方法中,使用两个或两个以上的预测模型对同一对象进行预测,然后对各个单独的预测结果适当地加权后取其平均值作为预测结果的预测模型。因此,各种预测方法组合权重的确定是一个重点、难点问题。要提高组合预测精度,必须研究合适的权重确定方法。

2.4.3 模型复杂性与模型选择

在预测过程中,模型的选择至关重要,是不是预测模型的形式越复杂、越完备,其预测精度就越高?如何根据已知数据选择合适的预测模型?人们经过长期研究认为,没有任何一种预测模型会在各种情况下都比其他方法或模型表现得更好。而正是由于没有任何一种预测模型的精度会肯定高于其他模型,因此提高预测模型的复杂程度不一定能提高预测精度。

事实上,备件预测模型都是建立在一定的假设条件(无限的过去都以现在为归宿,无限的未来都以现在为渊源)上的,同时利用参数估计的方法对历史数据进行“曲线拟合”,从而确定备件预测模型对未来的备件需求量进行预测。然而,假设条件不一定完全适合实际备件预测及其所处的环境。

备件预测模型完备性程度越高,所需考虑的影响备件需求量的因素就越多,就要求各因素间的相互影响的定量关系越清楚。事实上,以上这些都是难以达到的,所以所能做出的模型会在一定程度上影响备件预测的精度。

对于备件预测模型的选择,主要考虑预测精度和适用性。预测精度随着时间范围、数据类型(年份数据、季度数据等)和精度测定方法三个条件的不同而不同。

针对预测模型的适用性选择备件预测模型,可以提高备件预测的准确性,而备件预测模型的适用性主要由以下几个方面决定:

(1) 备件需求数据的特征,如数据的多少、随机性,以及表现季节性和趋势等。

(2) 备件需求预测的期限,是指进行短期预测、中期预测还是长期预测。

(3) 与备件需求量有关的环境的影响。

(4) 备件预测者的技巧。

同时,备件预测模型的选择还要考虑系统使用人员因素,即备件预测模型是否被决策者所接受。

目前用于需求规划、预测与决策的技术很多,主要有时间序列预测模型、一般解析法、回归分析法、智能算法(如神经网络、支持向量机、遗传算法等)、灰色预测与决策、马尔可夫预测、蒙特卡罗方法、多目标优化决策,以及它们之间的相互结合等。预测对象不同,预测目标和预测特点千差万别,因此,不可能有一种可适用于任何情况的统一的建模方法,而只能是在对具体情况做具体分析的基础上,根据已有的知识和经验,选择较为合适的预测方法,建立较为合适的预测模型。

第 3 章

武器装备备件品种确定方法



武器装备备件品种的确定是备件需求量预测的必要前提，若品种不能准确确定，需求量预测也就无从谈起。从军事效益来看，如果备件品种确定不准，必然造成某些部件损坏时无法进行更换修理，影响武器装备的可用性；从经济效益来看，如果备件品种过多，则必然占用过多的资金和储存空间，会造成浪费。在实际应用中，并不是对武器装备使用维修过程中的所有元器件、模板、组件等都进行储备，这就要求选用合理有效的方法来确定和预测备件品种。

3.1 概述

目前，确定武器装备备件品种的方法主要有专家打分法、层次分析法、模糊综合评判法、证据理论等。这些方法各具特点，也有各自的适用范围，既可单独使用，也可以结合使用。

1. 专家打分法

构成武器装备的零、组、部件成千上万，但深究其分类，可以区分为结构件、标准件和成品三大类。从使用与维修及其故障与消耗看，这三类之间既有相同点又有不同点。产品的故障及失效与其本身特性、安装特性、环境特性、使用及维修特性密切相关，只要分别对结构件、标准件和成品从不同角度提出导致其需要备件的各种因素，通过一些在装备研制中积累了丰富经验的专家对各种因素确认

和打分，对具体的得分值进行相差技术处理，即可定量确定备件需求，这种方法就是专家打分法。

2. 层次分析法

层次分析法是美国运筹学家、匹兹堡大学教授萨蒂（T.L.Saaty）于 20 世纪 70 年代初，为美国国防部研究“根据各个工业部门对国家福利的贡献大小而进行电力分配”课题时，应用网络系统理论和多目标综合评价方法，提出的一种层次权重决策分析方法。它是一种定性分析和定量分析相结合的评估决策方法，它将评估者对复杂系统的评估思维过程数学化。其基本思路是评估者通过将复杂问题分解为若干层次和若干指标，并在同一层次的各指标之间简单地进行比较、判断和计算，得出不同指标的重要度，为正确决策提供依据。其具体原理和步骤如下：

- （1）建立层次结构。建立层次结构即建立待评估指标的层次结构。
- （2）构造判断矩阵。在各层指标中进行两两比较，并引入判断尺度将其量化，构成判断矩阵。构造判断矩阵时，评估者要反复回答：两个指标 A_i 和 A_j 哪一个更重要，重要多少，需要对重要多少赋予一定数值，采用 1~9 比例尺度（标度），具体如表 3-1 所示。

表 3-1 1~9 标度表

标 度	含 义
1	表示两个指标比较，具有同样的重要性
3	表示两个指标比较，一个指标比另一个指标稍微重要
5	表示两个指标比较，一个指标比另一个指标重要
7	表示两个指标比较，一个指标比另一个指标重要得多
9	表示两个指标比较，一个指标比另一个指标极重要
2、4、6、8 是介于上述两个相邻判断的中值	
若指标 A_i 与 A_j 比较得 a_{ij} ，则 A_j 与 A_i 比较得 $1/a_{ij}$	

决策者进行两两指标之间重要程度的比较，可得如下结果：

	A_1	A_2	...	A_n
A_1	a_{11}	a_{12}	...	a_{1n}
A_2	a_{21}	a_{22}	...	a_{2n}
\vdots	\vdots	\vdots	\ddots	\vdots
A_n	a_{n1}	a_{n2}	...	a_{nn}

根据此结果，得到比较矩阵 $A = [a_{ij}]_{n \times m}$ 。

A 矩阵具有如下性质：

$$\begin{cases} a_{ij} > 0 \\ a_{ii} = 1 \\ a_{ij} = 1/a_{ji} \end{cases} \quad (3-1)$$

(3) 计算指标相对权重。

求解判断矩阵 A 的特征根：

$$AW = \lambda_{\max} W \quad (3-2)$$

式中： λ_{\max} 为判断矩阵 A 的最大特征根； W 为 λ_{\max} 所对应的特征向量，即权重向量。

(4) 一致性检验。由于客观事物的复杂性、主体认识的局限性及主体之间认识的多样性，判断矩阵不可能具有完全一致性，判断矩阵的不一致性会影响结果的准确性，因此在求权重之前，必须对判断矩阵 A 进行一致性检验，当不满足一致性要求时，要进行调整。对于 AHP 互反判断矩阵的一致性检验已有较多研究成果，Saaty 及其同事们的理论研究和社会实践，总结出了一致性检验的方法步骤：

① 计算一致性指标：

$$C.I. = \frac{\lambda_{\max} - n}{n - 1} \quad (3-3)$$

式中： n 为矩阵 A 的阶数。

② 求出随机一致性指标：

$$C.R. = \frac{C.I.}{R.I.} \quad (3-4)$$

随机一致性指标可采用表 3-2 所列的数值。

表 3-2 随机一致性指标

矩阵阶数	1	2	3	4	5	6	7	8	9	10
R.I.	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49

③ 判断矩阵的一致性检验标准：

当随机一致性指标 $C.R. \leq 0.1$ 或 $C.I. \leq 0.1 R.I.$ 时，一般认为判断矩阵的一致性是可以接受的，可利用上述方法求权重；否则，必须重新调整判断矩阵 A 中的元素值，再重新计算，直到满足一致性要求为止。

层次分析法能将人们的思维过程数学化、系统化,便于人们接受。该方法所需定量数据信息较少,在处理由相互联系、相互制约的众多因素构成的复杂而又往往缺乏定量数据的系统时有其特有的优势。目前它在复杂系统中的权重系数确定及方案优选中得到广泛应用。

3. 模糊综合评判法

模糊综合评判法(Fuzzy Comprehensive Evaluation, FCE),又称为模糊综合评估法、模糊综合决策法或模糊多元决策法,是解决定性和定量问题的经典方法,该方法在决策、预测、控制、评估等许多领域得到了广泛应用。建立模糊综合评判法的一级模型可归纳为以下五个步骤:

(1) 建立评估对象的指标论域 U :

$$U = \{u_1, u_2, \dots, u_n\}$$

即确定评估指标体系,解决了从哪些方面和用哪些指标来评估客观对象的问题。

(2) 确定评语等级论域 V :

$$V = \{v_1, v_2, \dots, v_m\}$$

即使得模糊综合评判得到一个模糊评估向量,被评估对象对各评语等级隶属程度的信息通过这个模糊向量表示出来,体现评估的模糊特性。

从技术处理来看,评语等级个数 m 通常要大于 4 而不超过 9。一方面, m 太大会超出人的语义区分能力,不易判断对象的等级归属;另一方面, m 太小又不符合模糊综合评判的质量要求。因此, m 过多或过少都对评估结果有不良影响,应以适中为宜。另外, m 取奇数者较多,这样除中间项之外,评语是对称的,如很好、好、一般、不好、很不好等。这样处理便于进一步计算评语隶属度。

(3) 进行单指标评估,建立模糊关系矩阵:

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nm} \end{pmatrix} \quad (3-5)$$

式中: r_{ij} 为 U 中指标 u_i 对应 V 中等级 v_j 的隶属关系,即相对于评估对象来说指标 u_i 被评为 v_j 等级的可能性,其实是一种单指标评估,它是模糊综合评判法的基础。

(4) 确定指标权向量 W 。 W 代表评估指标在被评估对象中的相对重要程度,

它在综合评估中用来对 R 做加权处理。通常 $W = \{w_1, w_2, \dots, w_n\}$, 且满足 $\sum_{i=1}^n w_i = 1$, $w_i \geq 0$ ($i=1, 2, \dots, n$)。

(5) 选择合成算子, 进行综合评估。模糊综合评判法的基本模型用公式表示为

$$B = W \circ R \quad (3-6)$$

式中: “ \circ ” 代表合成算子, 较常用的有加权平均算子等。

模糊综合评判法是在指标集和评语集之间建立模糊变换而产生的, 其特点是从广泛的调查中获得大量难以用确定的量来衡量的信息, 即模糊信息, 此方法只能用于模糊信息系统; 综合评判结果的可靠性和准确性依赖于合理选取指标和指标的权重分配。通过模糊评估, 能获得系统各替代方案优先顺序的有关信息, 且模糊的方法更接近于东方人的思维习惯和描述方法。因此, 它适用于评估层次结构复杂、不确定且信息不充分的社会经济系统及工程技术问题, 特别适用于对主观或定性指标进行评估。

4. 证据理论

证据理论是由 A.P.Dempster 首先提出, 并由 G.Shafer 进一步发展起来的用于处理不确定性问题的一种理论, 又称为 D-S 理论, 简称 DST。D-S 理论可以处理由“不知道”引起的不确定性, 并且不必事先给出知识的先验概率, 具有很大的灵活性。下面简要介绍该理论的一些基本知识。

1) 基本概念

定义 3.1 辨识框架: 证据理论中的每个概念和函数都是基于辨识框架的。若采用 Ω 表示样本空间, 其描述变量 X 的所有可能取值的非空、有限的穷举集合, 且所有在 Ω 内的元素是互补相容的, 则 Ω 称为所研究问题的辨识框架。辨识框架中的元素是相互排斥的, 对问题的描述也是完备的。

定义 3.2 命题: 它是辨识框架 Ω 的一个划分, 即 Ω 的子集。 Ω 中所有可能子集的集合用幂集 2^Ω 来表示, 它表示了所有可能的命题集。若 Ω 中有 N 个元素, 则幂集 2^Ω 中有 2^N 个元素, 每个幂集的元素对应于一个关于取值情况的子集(命题)。

定义 3.3 基本概率赋值 (Basic Probability Assignment, BPA) 函数: 一个证据可以为一个或者多个命题提供支持, 这种支持度可以用 BPA 函数表示。BPA 函数是 $[0,1]$ 范围的正数, 与支持某个假定的证据相联系。对任一个属于辨识框架 Ω 的子集 A (命题), 若函数 $m: 2^\Omega \rightarrow [0,1]$ 满足下列条件:

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{A \in \Omega} m(A) = 1 \end{cases} \quad (3-7)$$

则 m 称为辨识框架 Ω 上的 BPA 函数; $m(A)$ 为 A 的基本概率赋值或基本可信度, 它表示对命题的精确信任程度, 表示了对 A 的直接支持; \emptyset 为空集。第一个条件反映了对不可能事件(空命题)不产生任何信度, 即基本概率赋值为 0; 第二个条件说明了幂集 2^Ω 中全部元素的基本可信度分配值之和为 1。

2) 组合规则

定理 3.1 两个信任函数的组合规则(Dempster 组合规则): 设 Bel_1 和 Bel_2 是同一辨识框架上的两个信度函数, m_1 和 m_2 分别是其对应的基本可信度分配, 若 $A \subseteq \Omega$ 且 $m(A) > 0$, 则 A_i 和 B_j 为焦元, 焦元分别为 A_1, A_2, \dots, A_k 和 B_1, B_2, \dots, B_s 。

设

$$K = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j) < 1 \quad (3-8)$$

那么, 由式(3-9)定义的函数 $m: 2^\Omega \rightarrow [0, 1]$ 是基本可信度分配:

$$m(A) = \begin{cases} 0 & (A = \emptyset) \\ \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - K} & (A \neq \emptyset) \end{cases} \quad (3-9)$$

这称为 Dempster 组合规则。

式(3-9)中, $i = 1, 2, \dots, k$, $j = 1, 2, \dots, s$, K 反映了证据间的冲突程度, 其值越大说明证据之间的冲突越大。当 $K \neq 1$ 时, 正交和 m 也是一个基本可信度分配函数; 当 $K = 1$ 时, 则不存在正交和 m , 称 m_1 和 m_2 矛盾。分母 $1 - K$ 是归一化因子, 它的作用是避免在合成时将非 0 的信任赋给空集, 保证把融合的基本可信度分配规划到 $[0, 1]$ 。

定理 3.2 多个信任函数的组合规则: 设 $\text{Bel}_1, \text{Bel}_2, \dots, \text{Bel}_n$ 是同一辨识框架上的信度函数, m_1, m_2, \dots, m_n 分别是对应的基本可信度分配。如果 $\text{Bel}_1 \oplus \dots \oplus \text{Bel}_n$ 存在, 则 n 个信度函数的组合为 $((\text{Bel}_1 \oplus \text{Bel}_2) \oplus \text{Bel}_3) \oplus \dots \oplus \text{Bel}_n$, 由组合证据获得的最终证据与其组合次序无关。多个证据组合的计算可以用两个证据组合的计算递推得到。

3) 证据冲突的处理

设 α ($0 < \alpha < 1$) 为允许的冲突水平, 其值在证据融合之前事先设定。若 $K > \alpha$, 则可通过修改 Dempster 组合规则或者修改证据源模型(Dempster 组合规则不变)来减小冲突系数值, 使其满足允许的冲突水平。

证据理论不仅强调证据的客观性，而且重视证据估计的主观性，可以根据各种资料对系统各个部分状态的概率进行归纳与估计，并做出正确的决策。它具有比较强的理论基础，既能处理随机性所导致的不确定性，又能处理模糊性所导致的不确定性；采用区间估计而不是点估计，因此证据理论可以依靠证据的积累，不断地缩小假设集，从而增强对目标的认识，减少认识中存在的不确定性；能将“不知道”和“不确定”区分开来。把“不确定”的因素用信任函数或似然函数表示，通过融合减到最小，而对“不知道”的信息用分配到辨识框架的基本概率赋值函数表示；在多个不确定证据融合时，合成法则能够保证结果不受组合次序的影响，具有较强的容错能力，化解矛盾，在归一化过程中得到一致的结果。

3.2 基于相似分析法的备件品种确定方法

3.2.1 相似分析法简介

1. 整体决策过程规划

对于武器装备系统而言，由于同类装备发展的同源性和继承性，其任务相近，工作原理也相似。因此，备件的相似性是普遍存在的。相似分析法是根据新旧装备部分备件消耗存在某种程度的相似性，通过比较、分析、修正，对新旧备件进行相似度分析，判断其是否相似，从而进行备件的品种决策。相似分析法决策过程如图 3-1 所示。

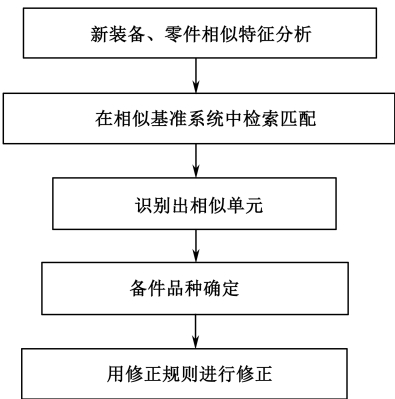


图 3-1 相似分析法决策过程

2. 相似理论简介

1) 相似理论进行问题研究的一般步骤

- (1) 通过对原系统的分析, 根据系统内部因素所满足的规律, 寻找相似元;
- (2) 分析相似元, 确定相似元的特征因素;
- (3) 量化特征因素;
- (4) 计算相似度。

2) 相似系统的几个基本概念

(1) 相似元: 系统间具有共同属性和特征, 而在数值上存在差异的组成要素构成的单元。当系统 A 中要素 a_i 与系统 B 中要素 b_i 为对应的相似要素时, 用相似元描述, 记为 $u_{ji}(a_i, b_i)$ 。如果系统间有 n 个相似元, 用 u_1, u_2, \dots, u_n 表示, 记为 $U = (u_1, u_2, \dots, u_n)$, 而每个相似元又包含若干特征要素, 如 u_i 中包含 $u_{i1}, u_{i2}, \dots, u_{ik}$ 。除了一般的特征要素 (具有确定的特征单值量) 外, 很多特征要素都存在不确定性因素, 即要素之间的属性和特征是模糊的, 如装备的结构式样和使用环境等, 因此, 相似元 $u_{ji}(a_i, b_i)$ 实际上是由多个一般要素和模糊要素组成的集合。

(2) 相似系统: 当两个系统间存在相似性要素和相似特性时, 则称两个系统为相似系统。显然, 两个系统间相似元越多, 相似元的值越大, 两个系统的相似程度就越高。相似系统可以看作相似元的集合。

(3) 系统相似度: 设系统 A 由 k 个要素组成, 系统 B 由 l 个要素组成, 相似要素为 n 个, $q(u_i)$ 表示相似元的值, β_i 为每个相似元的权重系数, 则系统间的相似度定义为

$$Q(A, B) = \frac{n}{k+l-n} \sum_{i=1}^n \beta_i q(u_i) \quad (3-10)$$

式中: $\frac{n}{k+l-n}$ 体现了系统间相似元数量的多少对系统相似度的影响; $\beta_i q(u_i)$ 体现了系统间相似元值的大小及其权重系数对系统相似度的影响。

3. 相似元的确定与优化

在研究新系统和备选系统的相似程度时, 如果选取的相似元不同, 就会选择不同的相似设备或部件, 很可能确定的最相似系统就不同。因此, 为了进行相对统一的比较, 必须对新系统和备选系统的相似元, 特别是其中的模糊要素进行量化。在此, 本书对系统设计特性、使用环境、可靠性维修性指标及功能特性等方面进行考虑, 选出相似要素, 构成相似元。系统相似度取决于各相似元的选取和

其相似程度的大小。

考虑新系统和备选系统间某相似元 $u(a_i, b_i)$ ，设该相似元中有 m 个特征要素，记为 s_1, s_2, \dots, s_m 。记 $u_j(a_i)$ 为 a_i 相对于要素 s_j 的特征值（具有确定的特征值向量），记 $u_j(b_i)$ 为 b_i 相对于要素 s_j 的特征值，则 a_i 、 b_i 相对于特征值 s_j 的特征比例系数记为

$$r_{ij} = \frac{\min\{u_j(a_i), u_j(b_i)\}}{\max\{u_j(a_i), u_j(b_i)\}} \quad (i=1, 2, \dots, n; j=1, 2, \dots, m; 0 \leq r_{ij} \leq 1) \quad (3-11)$$

求得相似元中各相似要素的特征值比例系数 r_{ij} 和隶属度 u_{ij} 后，还需要确定各特征要素的影响程度，即需要确定各特征要素的权重 d_j 。在权重不随相似特征本身变化而变化的情况下，可以采用专家打分法来确定各特征要素对相似元值的影响。对于特征要素权重的计算过程这里不再详述。得到各特征要素的权重后就可以利用

$$q(u_i) = \sum_{j=1}^m d_j r_{ij} \quad (3-12)$$

计算一般要素和模糊要素的相似元值，进而度量系统相似度大小。

4. 计算相似度并进行最终决策

相似系统的相似程度既要看系统间相似元的多少，也要考虑每对相似元相似程度的大小，因此系统相似度是相似元数量、相似元数值，以及每个相似元对系统相似度影响程度（权重）等因素的函数。除了相似要素数量、相似元值之外，每个相似元的权重也是影响系统相似度的一项主要指标，而不同相似元对系统相似度影响程度是不同的，各相似元的权重是否合理直接影响备选系统的有效性。相似元权重的确定可以参考层次分析法，以相似元作为评价因素，建立 AHP 层次结构，构造判断矩阵，通过计算可确定各个相似元的权重，得到权重值后就可以按式 (3-12) 计算系统相似度。

值得注意的是，该相似系统法既要满足最大值原则又要满足阈值，即在得到最相似系统后，也需要视相似度的大小来合理确定是否可以参照原系统进行备件品种配置。一般当 $q(u_i) \leq 0.65$ 时，就考虑重新选择相似系统。当达到相似度要求后，就可根据相似系统或单元进行决策。

3.2.2 示例分析

以某新型舰上短波综合通信系统为例，通过比较原有型号舰上短波综合通信

系统相似度的大小，来判断配置备件品种的情况。

1. 确定相似元和特征要素

根据舰上短波综合通信系统的使用与维修特点，选择本身构造 u_1 、使用特点 u_2 和维修方案 u_3 三个方面进行相似性分析，并将其再细分为各个特征要素，即

u_1 的特征要素={结构式样，故障率}； u_2 的特征要素={工作环境条件，工作强度，人员操作能力}； u_3 的特征要素={维修方式，维修能力}

2. 相似度分析与计算

将原有型号舰上短波综合通信系统与新型舰上短波综合通信系统进行比较，如表 3-3 所示。

表 3-3 新系统和原系统备选方案的相似元、特征要素及相应特征值

相似元	特征要素	新型舰上短波综合通信系统特征值	原有型号舰上短波综合通信系统特征值
本身构造	结构式样	A	A_1
	故障率	0.004	0.003
使用特点	工作环境条件	0.004	0.004
	工作强度	B	B_1
	人员操作能力	0.861	0.784
维修方案	维修方式	C	C_1
	维修能力	0.946	0.785

表 3-3 中，特征要素“人员操作能力”和“维修能力”用业务熟练人员占总人员的比例来表示。因此，对于新的或现有车体，特征要素“故障率”“人员操作能力”“维修能力”有确定的特征单值量，可以用式 (3-11) 来计算特征值比例系数 r_{ij} ；而其他特征要素为模糊要素，这里通过专家打分法进行赋值，完全相同时值为 1。

通过计算，得到原有型号舰上短波综合通信系统特征要素的特征值比例系数和相似度分值，如表 3-4 所示。同时，表 3-4 还给出了各相似元和各特征要素的权重系数，即可计算出各原有型号舰上短波综合通信系统与新型舰上短波综合通信系统的相似度。从表 3-4 可以看出，相似度 $q(u_i) \geq 0.65$ ，因此可以参照原有型号舰上短波综合通信系统进行器材品种确定。

表 3-4 原有型号舰上短波综合通信系统与新型舰上短波综合通信系统的相似度计算

相似元	特征要素	原有型号舰上短波综合通信系统相对于新型舰上短波综合通信系统的 r_{ij}
本身构造	结构式样 (0.4)	0.67
	故障率 (0.6)	0.75
使用特点	工作环境条件 (0.3)	1.00
	工作强度 (0.4)	0.83
	人员操作能力 (0.3)	0.91
维修方案	维修方式 (0.6)	0.95
	维修能力 (0.4)	0.83
相似度		0.71

3.3 基于模糊层次综合评判法的备件品种确定方法

3.3.1 模糊层次综合评判法简介

层次分析法是一种解决多目标的复杂问题的定性与定量相结合的决策分析方法。该方法将定量分析与定性分析结合起来,用决策者的经验判断各衡量目标能否实现标准之间的相对重要程度,并合理地给出每个决策方案的每个标准的权数,利用权数求出各方案的优劣次序,可比较有效地应用于那些难以用定量方法解决的问题。

模糊综合评判法是利用模糊集理论进行评价的一种方法。具体地说,该方法是应用模糊关系合成的原理,从多个因素对被评判事物隶属等级状况进行综合性评判的一种方法。模糊综合评判法不仅可对评价对象按综合分值的大小进行评价和排序,而且可根据模糊评价集上的值按最大隶属度原则去评定对象所属的等级。然而,模糊综合评判法不能解决评价指标间相关造成的评价信息重复问题,隶属函数的确定还没有系统的方法,而且合成的算法也有待进一步探讨,它是一种基于主观信息的综合评价方法。

模糊层次综合评判法是一种定性分析与定量分析相结合的、便于实际应用的系统分析、决策与预测方法。一般是先用层次分析法确定因素集,然后用模糊综合评判法确定评判效果。它综合了层次分析法和模糊综合评判法的优点,省去了层次分析法的繁杂计算,使之对复杂问题的决策条理化、层次化、定量化和模型化。这种模型组合技术,使层次结构模型简单、形象、清楚,因此广泛地适用于

非线性、多变量及动态复杂的大系统的识别、诊断、排序优化和最优控制等，尤其适用于宏观与微观的决策与预测。模糊层次综合评判法是在层次分析法之上，两者相互兼容，对评价有着很好的可靠性。

3.3.2 模型建立

模糊层次综合评判法首先应确定评语集，然后进行评价因素（主要有关键性、耗损性、经济性等）的判定进而进行权重分配，最后通过确定出的模糊矩阵进行评判。

1. 明确问题

首先通过对武器装备及其备件保障情况进行分析，初步确定影响备件品种确定的因素，然后选择一批既有丰富实践经验又有一定理论水平的专家，进行反复咨询、讨论和诊断分析，完成以下工作：

（1）确定目标 G 。在提高武器装备的战备完好性的同时对备件进行最优控制，以此确定备件品种即决策的目标层。

（2）确定评语集矩阵。为了准确地确定某种备件是否需要储备，对评价对象的评语集合选择为

$$U = \{u_1, u_2, u_3\} = \{\text{需储备, 可储备, 不需储备}\}$$

（3）设计评价指标体系 $T_i (i=1, 2, 3)$ 。评价指标体系的设计原则是，必须抓住备件的关键性、损耗性、经济性这三个方面的关键指标来全面而系统地设计指标体系。

T_1 ：关键性，即该件在装备系统中所起的作用及对系统性能影响程度的大小。通常分为关键件、重要件和一般件。越关键的件，它的失效对整个装备的影响越大，就越需要备件。

T_2 ：损耗性，即部件的损耗程度的大小。它与部件的固有可靠性、人为因素、工作环境及条件因素有关。越易损耗的部件，就越需要备件；不易损耗的部件，则不太需要备件。

T_3 ：经济性，即部件的成本、价格的高低。部件昂贵，则不仅其造价较高，而且其包装运输、维修保养的费用也往往较高，储备此类备件的效益较低。相反，价格低廉的部件，可以多储备一些，不会产生过多的经济浪费。所以，部件经济性在某种程度上也影响着备件品种的确定。

（4）建立评价指标集：

$$T = \{T_1, T_2, T_3\} = \{\text{关键性, 损耗性, 经济性}\}$$

2. 建立层次分析结构模型

依据层次分析结构建立模型,可以使对复杂问题进一步条理化、层次化与系统化。层次分析结构模型如图 3-2 所示。

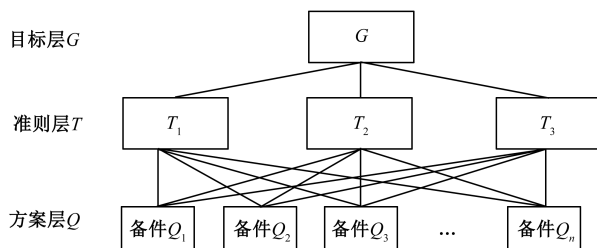


图 3-2 层次分析结构模型

3. 确定指标集权重

权重分配的方法很多,这里我们采用“群组评判法”,即请 N 个本行专家,将三个因素按其重要度排序,例如:

Z_1	T_1	T_3	T_2
Z_2	T_2	T_1	T_3
\vdots	\vdots	\vdots	\vdots
Z_k	T_1	T_2	T_3
\vdots	\vdots	\vdots	\vdots
Zh_N	T_2	T_3	T_1

其中: Z_k 为第 k 个专家, $k=1,2,\dots,N$ 。

若设 $Bk(u)$ 表示在第 k 行中排在 u 后面的因素数加上 1 的和,称为 u 的 Borda 数,则 $B(u)$ 为

$$B(u) = \sum_{i=1}^N Bk(u)$$

设 $Y = B(u_1) + B(u_2) + B(u_3)$, 从而可得各因素的权重度分别为

$$\alpha_1 = B(u_1)/Y$$

$$\alpha_2 = B(u_2)/Y$$

$$\alpha_3 = B(u_3)/Y$$

则权重集 $A = \{\alpha_1, \alpha_2, \alpha_3\}$ 。

4. 确定评价等级矩阵

首先将各制约因素对备件携带的影响程度按“需储备、可储备、不需储备”三个档次来区分，然后按通用的 1~5 标度法，将其量化得到评价矩阵 $P = \{P_1, P_2, P_3\}$ 。

5. 进行综合模糊评价

(1) 建立模糊评价表，如表 3-5 所示。

表 3-5 模糊评价表

部 件	α_i	U			专家人数
		需储备	可储备	不需储备	
Q_i	评价权重	P_1	P_2	P_3	N
Q_1	α_1	q_{11}	q_{12}	q_{13}	
Q_2	α_2	q_{21}	q_{22}	q_{23}	
Q_3	α_3	q_{31}	q_{32}	q_{33}	

(2) 建立模糊评价矩阵。

对部件 Q_j 建立模糊评价矩阵 SQ_j 。

$$SQ_j = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}$$

可转化为模糊频率矩阵 HQ_j 。

$$HQ_j = \begin{bmatrix} q_{11}/N & q_{12}/N & q_{13}/N \\ q_{21}/N & q_{22}/N & q_{23}/N \\ q_{31}/N & q_{32}/N & q_{33}/N \end{bmatrix}$$

(3) 求逻辑数 K 。逻辑数 K 依赖于单项指标 T_i 的可行度 C_i ：

$$C_i = \sum_{j=1}^3 \sum_{i=1}^3 P_i HQ_j \tag{3-13}$$

根据评价等级 3 是“可储备”的，也就是基本可行的，可以确定单项指标的临界可行度为 3。

显然，当 3 个 $C_i \geq 3$ 时， $K = 1$ ；否则， $K = 0$ 。

(4) 求综合评价值 WQ_i

$$WQ_i = K \times [A] \times [HQ_i] \times [P] \tag{3-14}$$

(5) 进行储备决策。确定储备可行性的临界指标 $WQ_i \geq 3$ ，即当 $WQ_i \geq 3$ 时，备件 Q_i 是需储备的。

3.3.3 示例分析

某型通用气源车主要用于向导弹包装箱或储运发射箱提供储存用低压干燥空气；向其他设备气瓶提供高压干燥空气（19MPa）；向导弹涡轮喷气发动机维护提供大流量气源。其中，手动截止阀是用于压缩机部位的备件之一，试采用模糊层次综合评价法综合评价该手动截止阀备件的储备必要性。

1. 明确问题

(1) 评语集矩阵。为了准确地确定某型武器装备的零部件是否需要储备备件，确定评价对象的评语集为

$$U = \{u_1, u_2, u_3\} = \{\text{需储备、可储备、不需储备}\}$$

(2) 评价因素。根据评价指标体系的设计原则，建立评价指标集为

$$T = \{T_1, T_2, T_3\} = \{\text{关键性, 损耗性, 经济性}\}$$

式中： T_1 表示备件的关键性； T_2 表示备件的损耗性； T_3 表示备件的经济性。

2. 权重分配

(1) 确定指标集权重。此外采用“群组评判法”，即邀请 10 个武器装备保障专家，对指标集的三个因素按其重要度进行排序：

Z_1	T_1	T_3	T_2
Z_2	T_1	T_2	T_3
Z_3	T_1	T_3	T_2
Z_4	T_1	T_2	T_3
Z_5	T_1	T_2	T_3
Z_6	T_1	T_2	T_3
Z_7	T_1	T_2	T_3
Z_8	T_1	T_2	T_3
Z_9	T_1	T_2	T_3
Z_{10}	T_1	T_2	T_3

利用公式 $B(u) = \sum_{k=1}^m B_k(u)$ 计算得

$$B(T_1) = 3 \times 10 = 30$$

$$B(T_2) = 2 \times 8 + 1 \times 2 = 18$$

$$B(T_3) = 2 \times 2 + 1 \times 8 = 12$$

$$Y = \sum B(T_i) = 30 + 18 + 12 = 60$$

从而可得各因素的权重度分别为

$$\alpha_1 = B(u_1)/Y = 30/60 = 0.5$$

$$\alpha_2 = B(u_2)/Y = 18/60 = 0.3$$

$$\alpha_3 = B(u_3)/Y = 12/60 = 0.2$$

故三种因素 T_1 、 T_2 、 T_3 对确定备件品种影响的权重集为

$$A = \{\alpha_1, \alpha_2, \alpha_3\} = \{0.5, 0.3, 0.2\}$$

(2) 确定评价等级矩阵 \mathbf{P} 。将各制约因素对携行备件的影响大小按“需储备、可储备、不需储备”三个档次量化,按通用的 1~5 标度法,将其量化得到评价矩阵 $\mathbf{P} = \{P_1, P_2, P_3\} = \{5, 3, 1\}$ 。

3. 综合评价

根据专家对该部件的评价表反馈情况,建立模糊评价表,如表 3-6 所列。

表 3-6 模糊评价表

部件	a_i 评价 权重	U			专家人数
Q_i		需储备	可储备	不需储备	
		5	3	1	
Q_1	0.5	8	2	0	10
Q_2	0.3	9	1	0	
Q_3	0.2	8	1	1	

(1) 对部件 Q_i 建立模糊评价矩阵 $\mathbf{S}Q_j$ 。

$$\mathbf{S}Q_j = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} = \begin{bmatrix} 8 & 2 & 0 \\ 9 & 1 & 0 \\ 8 & 1 & 1 \end{bmatrix}$$

(2) 可转化为模糊频率矩阵 $\mathbf{H}Q_j$ 。

$$\mathbf{H}Q_j = \begin{bmatrix} q_{11}/N & q_{12}/N & q_{13}/N \\ q_{21}/N & q_{22}/N & q_{23}/N \\ q_{31}/N & q_{32}/N & q_{33}/N \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.9 & 0.1 & 0 \\ 0.8 & 0.1 & 0.1 \end{bmatrix}$$

(3) 求逻辑数 K 。逻辑数 K 依赖于单项指标 T_i 的可行度 C_i 。

利用式 (3-13) 计算得

$$C_1 = 4.6, C_2 = 4.8, C_3 = 4.4$$

由于 3 个 C_i 均有 $C_i \geq 3$, 因此 $k = 1$ 。

(4) 求综合评价值 WQ_j 。

由式 (3-14) 计算得

$$\begin{aligned} WQ_j &= H \\ &= 1 \times (0.5, 0.3, 0.2) \times \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.9 & 0.1 & 0 \\ 0.8 & 0.1 & 0.1 \end{bmatrix} \times \begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix} = 4.8 \end{aligned}$$

(5) 进行储备决策。通过计算, 确定储备可行性的临界指标 $Q_i \geq 3$, 因此手动截止阀部件是需要储备的。

3.4 基于 D-S 理论和模糊综合评判的备件品种确定方法

由于影响备件品种确定的指标因素及配置效果本身具有复杂性和模糊性, 在充分吸收模糊综合评判和 AHP 优点的基础上, 提出基于 D-S 理论和模糊综合评判的备件品种确定方法。利用 D-S 理论处理不确定性问题的优势, 通过对专家评判意见和指标的两次证据融合, 增强评判结果的可区分性, 从而提高品种确定的科学性。

3.4.1 模型建立

建立基于 D-S 理论和模糊综合评判的备件品种确定模型的步骤如下:

- (1) 确定评语集, 即评判的等级, 并量化为评价值。
- (2) 明确评价指标, 分析影响备件品种确定的主要因素, 将其作为模型的评判指标。
- (3) 利用层次分析法计算指标的相对权重。
- (4) 利用 D-S 理论的组合规则, 确定备件配置效果在评语集上的可信度。

(5) 选择合适的算子，计算综合评判值。

(6) 方法对比，即从置信度和综合评判值两方面与基于 AHP 和模糊综合评判的品种确定方法进行比较。

(7) 根据综合评判值，确定备件品种。

其流程图如图 3-3 所示。

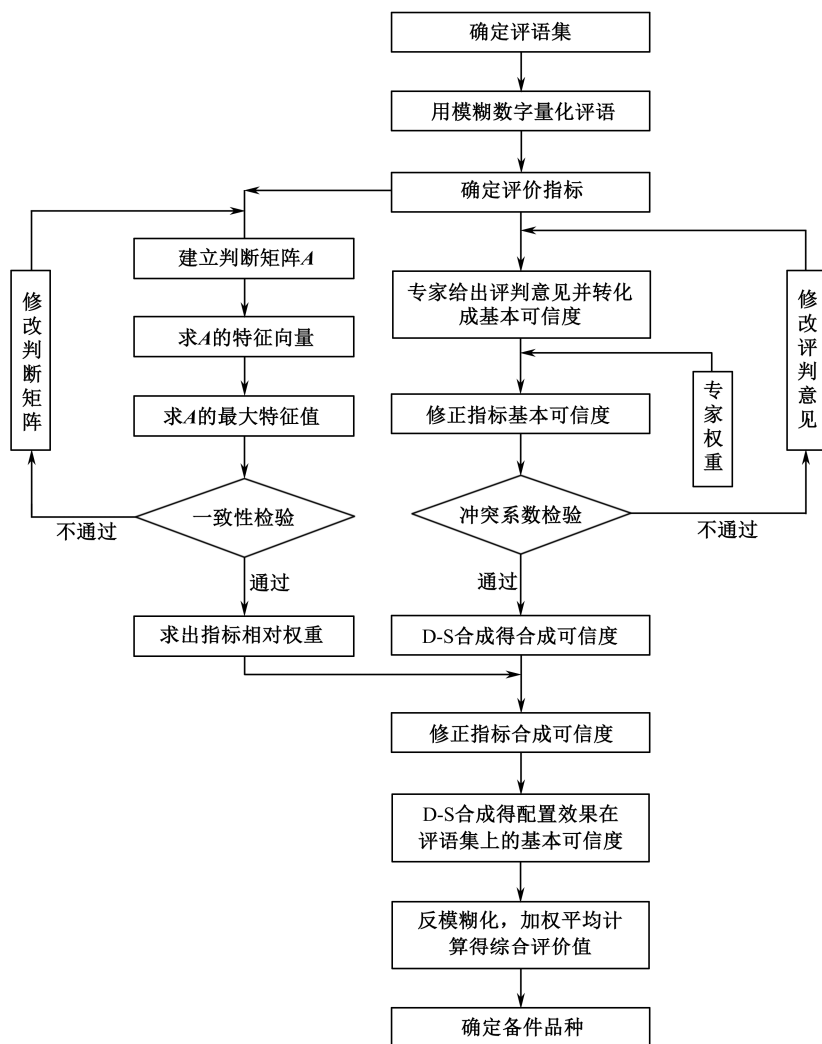


图 3-3 基于 D-S 理论和模糊综合评判的备件品种确定流程

3.4.2 实现步骤

1. 确定评语集

确定某单元是否需要备件,本质上是要看设置该类备件的效果如何。根据不同的效果评估等级可以有不同的品种确定策略,一般配置效果好的作为备件,效果差的则可不作为备件。本节的评语集给出 5 个等级,即“差”“较差”“一般”“较好”“好”,记为 $V_l(l=1,2,\dots,5)$,评语集 $V = \{V_1, V_2, V_3, V_4, V_5\}$,即证据理论中的辨识框架。记 $p(V) = \{p(V_1), p(V_2), p(V_3), p(V_4), p(V_5)\}$ 为评语集的模糊评价值,对应的取值分别为 $\{1, 2, 3, 4, 5\}$ 。

2. 明确评价指标

影响备件品种确定的因素很多,其中主要因素有:

(1) 关键性: 零部件在武器装备系统中所起的作用以及对系统性能影响程度的大小。零部件通常可根据单元故障的严酷度分为关键件、重要件和一般件,越关键的件,它的故障对系统的影响越大,就越需要备件或者说设置备件的效果越好。关键性对于备件品种的确定有着较大的决定作用。一般来讲,关键件对系统性能影响最大。其一旦损坏会导致武器装备系统失效,甚至会危及装备或人身安全。因此,这类件应主要考虑其军事效益,以确保备件的及时供应。在品种确定时,应尽量靠前(向基层)配置。而对于一般件,由于其损坏只会使武器装备局部的性能下降,不会对装备系统有严重影响,其供应允许有一定时间的延迟,因此,该类件以考虑经济效益为主,对于其中的贵重件应尽量靠后配置,对于重要件则应对其效果和费用进行综合权衡而定。

(2) 损耗性: 零部件损耗程度的大小,既包括自然因素又包括人为因素造成的损耗。它与零部件的固有可靠性、使用环境、操作及维修管理水平等因素有关。越易损耗的零部件,就越需要备件。损耗性的大小与备件品种确定密切相关。损耗性高的单元,其配置应尽量靠前;而损耗性低的单元,显然为提高其利用率应尽量向后配置。

(3) 易得性: 包括两方面的含义。其一,备件是否是标准件,如果是标准件,那么意味着有供应商生产,在市场上可以买到,易得性较好;如果是非标准件,供应商还没有确定,加工所需的技术资料也不齐全,则易得性较差。其二,备件的交货提前期是否很长,供应商在国外还是在国内,在外地还是在本地,供应商

有没有存货等。如交货期很短,供应商在本地且备有存货,说明该备件易得性较好,否则较差。易得性越差的零件就越有必要作为备件。

(4) 可更换性:可更换性决定了备件两个层面的属性。一方面与选择备件装配的最小组合单位有关,即是选择已装配好的一组零件作为一个基本备件单位,还是将其再细分成零件,以更小的零件单位为基本备件单位。如果这组零件装配要求特殊的装配工具和技术,而维修单位又不具备这些条件,往往只能把这一组装配好的组合件作为一个整体当作基本的备件单位。因为如果储备它的子零件,但维修单位没有技术条件把它装配起来,对维修工作就没有帮助。因此,选择越小的备件单位可更换性越差。另一方面与可更换备件的时机有关,而这往往与维修计划或维修级别有关。更换某一备件往往涉及不同程度的维修工作,如大修或小修。在某一维修级别储备的备件,应当是此维修级别能够成功地进行拆装更换的。对没有把握正确更换的备件,应当少储备或不储备。确定备件品种需要对指定的维修级别的能力有一个恰当的评价。任何一个维修级别,受客观条件和人员技术水平的制约不可能涵盖所有维修项目。因此,备件不应当盲目求全。由此可见,可更换性越好的零件越有作为备件的必要。

(5) 经济性:设置某备件成本、价格的大小,包括备件的购置费、储存保养费及储存损失费。导弹备件中,除少数光学、电子器件及专用油液的储存保养费及损失费较高外,其他备件则主要考虑购置费。越是贵重的零部件,不仅购置价格高,而且库存费用往往较高。而价格低廉的备件,即使多备几个也不会造成过大的浪费。因此,备件经济性的好坏也影响着备件品种和数量的确定。

上面这些影响因素很多不易量化,只能通过模糊化处理,而可以定量描述的其数据来源也并不可靠,它们与备件品种的关系是模糊的,再者设置备件效果的好坏本身也带有模糊性。为此,选择它们作为评判指标是合理的,指标层次结构如图 3-4 所示。

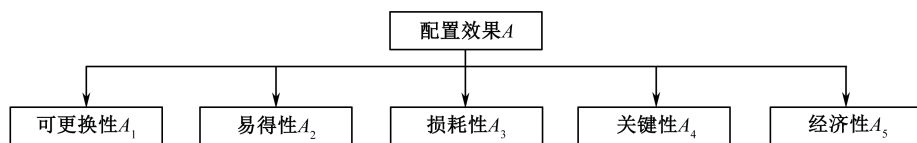


图 3-4 指标层次结构

3. 确定指标权重

利用图 3-4 的指标层次结构,采用层次分析法确定指标相对权重。

4. 确定配置效果在评语集上的可信度

1) 确定基本可信度分配

专家根据知识和经验给出对某指标定性的评判意见, 并进行定量描述, 再结合证据理论的相关原理将其转换为基本可信度分配。设有 n 位专家参与评价, 第 r 位专家对某项指标 A_i 的评价记为 $m_{A_i}^r(V_l)$, 则其基本可信度分配函数为

$$m_{A_i}^r(V_l) = \begin{cases} p_1 (A_i = \{V_1\}) \\ \vdots \\ p_l (A_i = \{V_l\}) \\ p_5 (A_i = \{V_5\}) \\ 0 \text{ (其他)} \end{cases} \quad (3-15)$$

式中: $\sum_{l=1}^5 p_l = 1$, p_l 为专家对指标 A_i 隶属等级的评价值。对于备件品种确定的各个指标, 依据专家的评判意见, 可得到相应的基本可信度分配。

2) 修正基本可信度分配

根据各位专家的权重, 修正各指标的基本可信度分配, 记为 $\beta_{A_i}^r(V_l)$ 。设 n 位专家的权重为 $\{w_1, w_2, \dots, w_n\}$, 且 $\sum_{r=1}^n w_r = 1$, 则专家对指标 A_i 的评判可靠度为

$$\eta_r = \varepsilon_r \frac{w_r}{\max_{1 \leq r \leq n} w_r} \quad (3-16)$$

式中: ε_r 为反映专家经验和偏好的系数, $0.9 \leq \varepsilon_r \leq 1$ 。

通过修正 $m_{A_i}^r(V_l)$ 求得 $\beta_{A_i}^r(V_l)$, 即

$$\beta_{A_i}^r(V_l) = \eta_r \times m_{A_i}^r(V_l) \quad (3-17)$$

3) 冲突系数检验

定义 $K = \sum_{\cap V_l = \emptyset} \prod_{1 \leq r \leq n} \beta_{A_i}^r(V_l)$ 为冲突系数, 设 α ($0 < \alpha < 1$) 为允许的冲突水平,

其值在专家评判前事先设定。若 $K < \alpha$, 则可利用 D-S 组合规则进行融合; 反之, 则表明专家的评判意见冲突较大, 不能直接进行证据融合。这里采取修改证据源模型的方式来减小冲突系数, 即让出现冲突的证据源 (相应专家) 重新给出评判意见, 形成新的基本可信度分配 $m_{A_i}^r(V_l)$, 再重复步骤 1) 和 2)。

4) 确定合成可信度

冲突系数检验若 $K < \alpha$, 则可利用 3.1 节的定理 1 对 $\beta_{A_i}^r(V_l)$ 进行证据组合,

确定指标的合成可信度 $m_{A_i}(V_l)$ 。

5) 修正合成可信度

采用 3.2.1 节的 AHP 计算各指标的权重 $\{w_{A_1}, w_{A_2}, w_{A_3}, w_{A_4}, w_{A_5}\}$, 且 $\sum_{i=1}^5 w_{A_i} = 1$ 。

指标 w_{A_i} 的相对重要度为

$$\eta_{A_i} = \varepsilon_{A_i} \frac{w_{A_i}}{\max_{1 \leq i \leq 4} w_{A_i}} \quad (3-18)$$

式中: ε_{A_i} 为指标权重修正系数, $0.9 \leq \varepsilon_{A_i} \leq 1$ 。

A_i 的合成可信度分配值修正为

$$\beta_{A_i}(V_l) = \eta_{A_i} \times m_{A_i}(V_l) \quad (3-19)$$

6) 确定评语集上的基本可信度

用 3.1 节的定理 3.1 对各指标的 $\beta_{A_i}(V_l)$ 进行证据组合, 得到备件配置效果在评语集上的基本可信度 $m(V_l)$, 并将其归一化。

5. 计算综合评判结果

反模糊化, 选用加权平均算子, 可计算出备件配置效果的综合评价值为

$$E = \sum_{l=1}^5 m(V_l) p(V_l) / \sum_{l=1}^5 m(V_l) \quad (3-20)$$

6. 备件品种确定

依据算得的综合评判值, 按照一定规则确定是否设置该零件为备件, 并定性描述其储备量。由于综合评价值范围是 $[1, 5]$, 在人们对事物判断的思维习惯以及征求有关专家意见的基础上, 确定取 3 为判断该件是否作为备件的临界值, 并得到以下策略: 当 $1 \leq E < 3$ 时, 该类零部件不宜列为备件品种; 当 $3 \leq E < 4$ 时, 该类零部件可列为备件品种, 在消耗量归整时, 取小的整数值; 当 $4 \leq E \leq 5$ 时, 该类零部件应列为备件品种, 在消耗量归整时, 取大的整数值。

3.4.3 示例分析

专家组有 4 人分别对各指标进行定性评判, 并转换成基本可信度分配。表 3-7 为根据各专家对指标 A_i 的评判意见而转换成的基本可信度分配。

表 3-7 A_1 的基本可信度分配

$m_{A_1}^r(V_l)$	V_1	V_2	V_3	V_4	V_5
$m_{A_1}^1(V_l)$	0.0	0.0	0.0	0.3	0.7
$m_{A_1}^2(V_l)$	0.0	0.0	0.2	0.2	0.6
$m_{A_1}^3(V_l)$	0.0	0.0	0.1	0.3	0.6
$m_{A_1}^4(V_l)$	0.0	0.0	0.0	0.2	0.8

设专家权重 $\{w_1, w_2, w_3, w_4\} = \{0.20, 0.35, 0.15, 0.30\}$ ，取 $\varepsilon_r = 0.95$ 。根据式 (3-18) 和式 (3-19) 求得 $\beta_{A_1}^r(V_l)$ ，表 3-8 为 A_1 修正后的基本可信度分配。

表 3-8 A_1 修正后的基本可信度分配

$\beta_{A_1}^r(V_l)$	V_1	V_2	V_3	V_4	V_5
$\beta_{A_1}^1(V_l)$	0.0000	0.0000	0.0000	0.1629	0.3801
$\beta_{A_1}^2(V_l)$	0.0000	0.0000	0.1900	0.1900	0.5700
$\beta_{A_1}^3(V_l)$	0.0000	0.0000	0.0407	0.1221	0.2442
$\beta_{A_1}^4(V_l)$	0.0000	0.0000	0.0000	0.1628	0.6512

取 $\alpha = 0.4$ ，经检验 $K < \alpha$ ，表明专家评判意见一致性较高，冲突系数检验通过，可以进行证据融合。由证据组合规则求得的 $m_{A_1}^{1234}(V_l)$ 为合成可信度 $m_{A_1}(V_l)$ ，结果见表 3-9。

表 3-9 A_1 的合成可信度

合成	V_1	V_2	V_3	V_4	V_5	K
$m_{A_1}^{12}(V_l)$	0.0000	0.0000	0.1187	0.1721	0.6780	0.2682
$m_{A_1}^{123}(V_l)$	0.0000	0.0000	0.0960	0.1592	0.7216	0.2029
$m_{A_1}^{1234}(V_l)$	0.0000	0.0000	0.0255	0.0846	0.8838	0.2993

依照同样的步骤和过程可以得到全部指标的合成可信度 $m_{A_i}(V_l)$ ，见表 3-10。

表 3-10 所有指标的合成可信度

$m_{A_i}(V_l)$	V_1	V_2	V_3	V_4	V_5
$m_{A_1}(V_l)$	0.0000	0.0000	0.0255	0.0846	0.8838
$m_{A_2}(V_l)$	0.0000	0.1556	0.1339	0.2757	0.4196
$m_{A_3}(V_l)$	0.0000	0.0132	0.0273	0.0419	0.8651
$m_{A_4}(V_l)$	0.0000	0.1313	0.1695	0.1993	0.4497
$m_{A_5}(V_l)$	0.0000	0.0000	0.2337	0.3161	0.4510

利用层次分析法求各指标的判断矩阵及相对权重，结果见表 3-11。

表 3-11 判断矩阵及相对权重

指标	A_1	A_2	A_3	A_4	A_5	w_i
A_1	1	1/3	1/2	1/2	1/2	0.0478
A_2	3	1	2	2	2	0.1349
A_3	5	2	1	1	1	0.2505
A_4	9	7	1	1	1	0.4756
A_5	2	1/2	1/3	1/4	1	0.0912

设 $\varepsilon_{A_i} = 0.9$ ，根据式 (3-18) 求得 $\eta_{A_i} = \{0.0905, 0.2553, 0.4740, 0.9000, 0.1726\}$ ，再按式 (3-19) 修正指标的合成可信度分配得 $\beta_{A_i}(V_l)$ ，结果见表 3-12。

表 3-12 修正后的合成可信度

$\beta_{A_i}(V_l)$	V_1	V_2	V_3	V_4	V_5
$\beta_{A_1}(V_l)$	0.0000	0.0000	0.0023	0.0077	0.0800
$\beta_{A_2}(V_l)$	0.0000	0.0397	0.0342	0.0704	0.1071
$\beta_{A_3}(V_l)$	0.0000	0.0063	0.0129	0.0199	0.4101
$\beta_{A_4}(V_l)$	0.0000	0.1182	0.1526	0.1794	0.4047
$\beta_{A_5}(V_l)$	0.0000	0.0000	0.0403	0.0546	0.0778

以评语集为辨识框架，以各指标为证据，再次利用证据组合规则对各指标的 $\beta_{A_i}(V_l)$ 进行合成，求得的 m_{12345} 即备件配置效果在评语集上的基本可信度 $m(V_l)$ ，并将其归一化，合成过程见表 3-13。

表 3-13 评语集上的基本可信度

合成	V_1	V_2	V_3	V_4	V_5
m_{12}	0.0000	0.0366	0.0334	0.0713	0.1682
m_{123}	0.0000	0.0265	0.0298	0.0584	0.4772
m_{1234}	0.0000	0.0783	0.1009	0.1307	0.6062
m_{12345}	0.0000	0.0720	0.1010	0.1331	0.6168
归一化	0.0000	0.0780	0.1094	0.1442	0.6683

通过式 (3-20) 求得综合评判值 $E=4.4029$ ，依照备件品种确定对策，该类零部件应该列为备件品种，在消耗量归整时，取大的整数值。

采用层次分析法计算得到的置信度 B 见表 3-14。

表 3-14 采用层次分析法的计算结果

AHP	V_1	V_2	V_3	V_4	V_5
B	0.0000	0.0906	0.1392	0.1816	0.5949

为便于比较，作直方比较图（图 3-5）。可以看出，经过证据融合得到的结果降低了评语 $\{V_2, V_3, V_4\}$ 上的置信度，增强了对评语 V_5 的置信度，很大程度上改善了评语等级的可区分性，有利于做出科学可靠的决策。

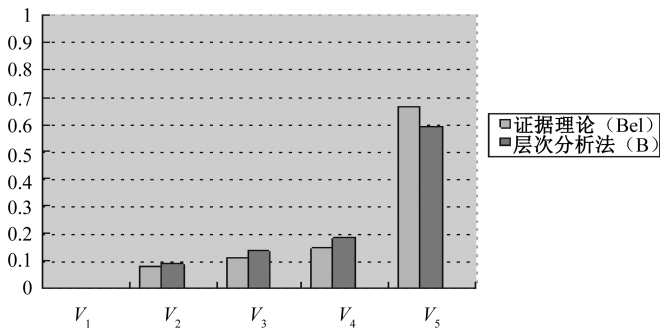


图 3-5 两种方法的结果比较

最后,应用 AHP 和模糊综合评判的备件品种确定方法进行计算,得到综合评判值 $e=4.2808$ 。尽管也能得到相同的备件品种决策,但是 $E > e$,该方法结果更优,更具有说服力。

3.5 基于灰色评估法的备件品种确定方法

备件品种确定从运筹学角度来看是典型的多目标决策问题,可采用逻辑决断法、价值工程法和可靠性等方法来进行决策,但由于这些方法本身固有的缺陷,在决策过程中无法很好地体现备件储备各指标间的相关性,指标间的合理综合等关键问题一直无法得到解决。此外,备件品种确定工作是一个复杂巨系统,涉及大量随机、不确定的因素。备件品种确定既有已知的白色信息,又有未知的黑色信息,更多的是一般定性了解的灰色信息,其系统实质就是一个灰色系统。因此,采用灰色评估法,能够综合考虑备件的各种属性特点、诸多人为因素,也能够吸收各类型底层因素的影响,将其逐层综合、归纳,最终得到对总评价目标的综合评价价值,使评估结果更加合理,符合客观实际。

3.5.1 评价指标体系的建立

通过对影响备件储备的相关因素进行分析,如备件的损耗性、关键性、经济性等,遵循科学性、全面性、导向性、综合性、可比性等基本原则,运用层次分析法构建备件品种决策指标体系,见表 3-15。其中,目标层 Z 为备件的储备效果,准则层为 $U_i (i=1,2,\dots,n)$, U 为准则层 U_i 所组成的集合, $U = \{U_1, U_2, \dots, U_n\}$, 指标层为 $u_{ij} (i=1,2,\dots,n; j=1,2,\dots,m)$ 。具体的指标说明如下。

1. 备件关键性 U_1

(1) 对装备影响程度 u_{11} : 关键性指标最重要的体现是对装备的影响程度,影响程度高,关键性越强,越应该优先考虑对其储备。

(2) 可替代性 u_{12} : 可替代性影响着备件的关键性,可替代性越强,备件损坏后对装备的影响程度越低,备件的重要程度降低,进行备件储备的作用降低;反之,备件储备的作用增加。

2. 备件损耗性 U_2

(1) 备件的可靠性 u_{21} : 备件的可靠性是备件的固有特性,备件的可靠性越

高，备件越不容易损坏，备件的损耗性越低。

(2) 使用环境 u_{22} ：备件的使用环境从另一方面影响备件的损耗，相同备件，使用环境越恶劣，备件的损耗越大；反之，备件的损耗越小。

(3) 备件管理水平 u_{23} ：备件的管理水平越高，由自然或人为因素造成的备件损耗越少，备件储备的必要性就越低，备件的管理水平越低，备件储备的必要性就越高。

3. 备件经济性 U_3

(1) 备件的购置费 u_{31} ：依据备件的购置费用的高低来判定备件的重要性。通常情况下，购置费用高的备件在储备时越要慎重，越不宜在下级仓库储备。

(2) 储存费 u_{32} ：备件储存费的高低反映了备件储备的难易程度，储存费越高，备件越难储存，备件储存的经济性越低。

4. 供应难易性 U_4

(1) 采购难易程度 u_{41} ：采购困难度高的备件一般要重点考虑储备，容易采购的备件可以少储备甚至不储备。

(2) 供应商数量 u_{42} ：通过对备件的供应市场的调查情况来判定，如果供应商数量增多，则供应难度就会降低。

(3) 供应能力 u_{43} ：备件的市场供应能力增强，则备件的供应难度降低。

表 3-15 备件储备品种决策指标体系

决策目标	一级指标	二级指标
备件周转储备效果 Z	备件关键性 U_1	对装备影响程度 u_{11}
		可替代性 u_{12}
	备件损耗性 U_2	备件的可靠性 u_{21}
		使用环境 u_{22}
		备件管理水平 u_{23}
	备件经济性 U_3	备件的购置费 u_{31}
		储存费 u_{32}
	供应难易性 U_4	采购难易程度 u_{41}
		供应商数量 u_{42}
		供应能力 u_{43}

3.5.2 评价模型的建立

灰色理论是从信息是否完备出发,将系统分为非灰色系统和灰色系统。其中:“黑”表示信息完全缺乏,“白”表示信息不充分、不完全;将信息不完全的元素称为灰元,信息不完全的系统称为灰色系统,信息不完全的关系称为灰关系。灰色聚类是根据“灰类”的灰数及白化权函数将一些观测指标或观测现象聚集成若干个可以定义类别的方法。灰色理论对信息不确定或不完全确定的小样本系统有明显的理论分析和实践操作优势。

为提高评价结果的真实可靠性,本项目选用灰色评估方法。灰色评估法是基于灰色系统的理论和方法,是针对预定的目标,通过系统分析,对某个系统或所属因子在某一时段所处的状态做出一种半定性半定量的评估与描述,以便在更高层次上对系统的综合效果与整体水平形成一个可供比较的概念与类别。

利用 Delphi 法和层次分析法拟定梯阶层次结构的决策(综合评价)指标体系,确定体系中各元素的权重,再利用灰色系统理论确定评估灰类,计算得出各个评价灰类的向量,最终得出评价结果,并根据结果做出决策。

1. 确定指标权重

在已构建决策指标体系的基础上,采用专家问卷调查法,利用 AHP 和灰色关联方法,对《权数评判专家咨询表》进行统计分析计算,求得二层指标的权重为 A_1, A_2, A_3 , 进而构造比较判断矩阵。

假设第一层指标 $U_i (i=1,2,\dots,n)$ 的权重为 $b_i (i=1,2,\dots,n)$, 则各指标的权重集为 $A=(b_1, b_2, \dots, b_n)$, 其中 $b_i \geq 0$, 且 $\sum_{i=1}^n b_i = 1$; 第二层指标 $u_{ij} (i=1,2,\dots,n; j=1,2,\dots,m)$ 的权重为 $b_{ij} (i=1,2,\dots,n; j=1,2,\dots,m)$, 其中 $b_{ij} \geq 0$, 且 $\sum_{j=1}^m b_{ij} = 1$ 。

2. 确定评价等级

为实现定性指标的量化,由多名专家组成的评价小组按照备件储备效果的好坏制定评价等级为 I、II、III、IV, 相应评语集 $E=\{\text{好, 较好, 一般, 差}\}$, 给每个评语赋予相应的分值, 得到评价等级向量 $G=(9,7,5,2)$, 见表 3-16。备件储备效果评价区间如下: 好, 8~10; 较好, 6~8; 一般, 3.5~6; 差, 0~3.5, 要求只对储备效果为良好以上(包括良好)的备件进行存储。

表 3-16 评价指标等级标准

等级标准	评分
好	9
较好	7
一般	5
差	2

3. 确定评价矩阵

假定共有 k 个专家参与评估, 则第 s 个专家对第 i 个指标的评分记为 d_{si} , 则 k 位专家对 n 个指标的评价分数构成评价样本矩阵为

$$\mathbf{D} = (d_{si}) = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ \vdots & \vdots & & \vdots \\ d_{s1} & d_{s2} & \cdots & d_{sn} \\ \vdots & \vdots & & \vdots \\ d_{k1} & d_{k2} & \cdots & d_{kn} \end{bmatrix} \quad (3-21)$$

4. 评价矩阵计算

1) 白化权函数的选择

根据需要, 本书选择最常用的三种白化权函数进行计算。

(1) 上端级。灰数为 $\otimes = [d_1, +\infty)$, 其函数形式为

$$f_1(d_{si}) = \begin{cases} d_{si}/d_1 & , d_{si} \in [0, d_1] \\ 1 & , d_{si} \in [d_1, +\infty] \\ 0 & , d_{si} \in [-\infty, 0] \end{cases} \quad (3-22)$$

(2) 中间级。灰数为 $\otimes = [0, d_1, 2d_1)$, 其函数形式为

$$f_2(d_{si}) = \begin{cases} d_{si}/d_1 & , d_{si} \in [0, d_1] \\ 2 - d_{si}/d_1 & , d_{si} \in [d_1, 2d_1] \\ 0 & , d_{si} \in [0, 2d_1] \end{cases} \quad (3-23)$$

(3) 下端级。灰数为 $\otimes = [0, d_1, d_2)$, 其函数形式为

$$f_3(d_{si}) = \begin{cases} 1 & , d_{si} \in [0, d_1] \\ \frac{d_2 - d_{si}}{d_2 - d_1} & , d_{si} \in [d_1, d_2] \\ 0 & , d_{si} \in [0, d_2] \end{cases} \quad (3-24)$$

其中,称白化权函数转折点 d_1 、 d_2 的值为阈值,该值由评价指标等级评分标准确定。

2) 计算灰色统计数及模糊权矩阵

利用灰色统计数法确定灰数的白化权函数,求 $f(d_{si})$,进一步得出判断矩阵的灰色统计数 n_{ij} 和总灰数统计数 n_i ,从而算出 k 位专家对第 i 个评价指标主张第 j 种评价等级的灰色权值 k_{ij} ,即

$$n_{ij} = \sum_{s=1}^k f_j(d_{si}) \quad (3-25)$$

$$n_i = \sum_{j=1}^n n_{ij} \quad (3-26)$$

$$k_{ij} = \frac{n_{ij}}{n_i} \quad (3-27)$$

最终构成单因素模糊权矩阵 R ,即

$$R = (k_{ij}) = \begin{bmatrix} k_{11} & \cdots & k_{1j} & \cdots & k_{1r} \\ \vdots & & \vdots & & \vdots \\ k_{i1} & \cdots & k_{ij} & \cdots & k_{ir} \\ \vdots & & \vdots & & \vdots \\ k_{n1} & \cdots & k_{nj} & \cdots & k_{nr} \end{bmatrix} \quad (3-28)$$

式中: r 为指标数。

5. 求解并决策

对模糊权矩阵和单因素模糊评判矩阵进行复合运算,得出模糊综合评判矩阵。对指标体系的第一层进行综合评价,得到 $B_i = A_i \times R_i$ 。再对 B_i 和 $G = (g_1, g_2, \cdots, g_m)$ (可由专家确定的评价等级向量得到)进行复合计算,求出综合评价的结果 $Z = BG^T$,并依据器材储备效果评价区间对评价结果进行决策。

3.5.3 示例分析

以某型舰面指挥通信设备系统的某型备件储备情况为例,选取8名专家对决策指标体系中的每个指标进行评分(规定专家打分的范围为1~10分),并采用灰色评价法对该备件的储备效果进行评分,最终决定是否对该种备件进行储备。

1. 确定指标权重

采用层次分析法和灰色关联方法确定各层级的指标权重。

第一层指标的权重向量为

$$A = (0.358, 0.286, 0.157, 0.199)$$

第二层指标的权重向量为

$$A_1 = (0.665, 0.335), A_2 = (0.368, 0.286, 0.346),$$

$$A_3 = (0.506, 0.494), A_4 = (0.436, 0.239, 0.325)$$

2. 确定指标评价样本矩阵

要求专家均按各指标对备件存储的效果的影响进行评分，见表 3-17。

表 3-17 指标评价样本矩阵

专家	指 标									
	u_{11}	u_{12}	u_{21}	u_{22}	u_{23}	u_{31}	u_{32}	u_{41}	u_{42}	u_{43}
k_1	9	9	7	8	5	6	8	5	6	4
k_2	10	9	6	9	1	6	6	3	5	6
k_3	8	9	7	8	5	6	7	2	6	6
k_4	9	8	8	8	3	5	7	5	5	5
k_5	7	10	6	8	3	6	6	4	5	3
k_6	9	9	6	8	2	6	8	4	6	5
k_7	9	9	6	9	2	7	8	4	5	5
k_8	8	9	8	7	3	5	8	3	5	5

3. 选择白化权函数

从实际需要出发，确定“好”“较好”“一般”“差”4个灰类，与其对应的白化权函数图见图 3-6~图 3-9。

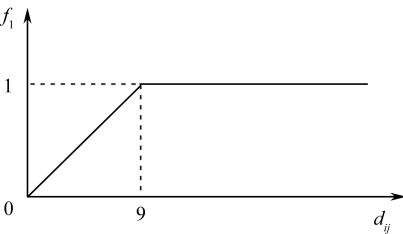


图 3-6 “好”类白化权函数

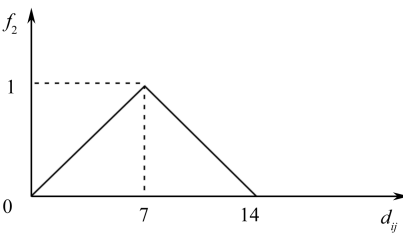


图 3-7 “较好”类白化权函数

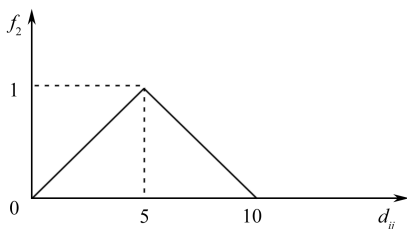


图 3-8 “一般”类白化权函数

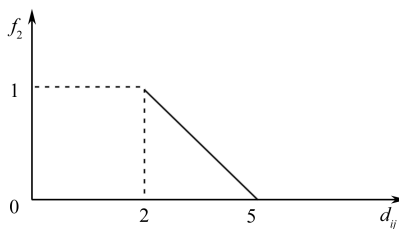


图 3-9 “差”类白化权函数

4. 计算灰色统计数和权矩阵

以第二层指标 u_{11} （对装备影响程度）为例，依据四类白化权函数的图形及表达式计算灰色统计数和权矩阵，并求出模糊综合评价矩阵。

(1) “好”类：

$$\begin{aligned} n_{11} &= f_1(d_{11} + d_{21} + d_{31} + d_{41} + d_{51} + d_{61} + d_{71} + d_{81}) \\ &= f_1^{(9)} + f_1^{(10)} + f_1^{(8)} + f_1^{(9)} + f_1^{(7)} + f_1^{(9)} + f_1^{(9)} + f_1^{(8)} \\ &= 1 + 1 + \frac{8}{9} + 1 + \frac{7}{9} + 1 + 1 + \frac{8}{9} = 7.556 \end{aligned}$$

(2) “较好”类：

$$\begin{aligned} n_{12} &= f_2(d_{11} + d_{21} + d_{31} + d_{41} + d_{51} + d_{61} + d_{71} + d_{81}) \\ &= f_2^{(9)} + f_2^{(10)} + f_2^{(8)} + f_2^{(9)} + f_2^{(7)} + f_2^{(9)} + f_2^{(9)} + f_2^{(8)} = 6.143 \end{aligned}$$

(3) “一般”类：

$$\begin{aligned} n_{13} &= f_3(d_{11} + d_{21} + d_{31} + d_{41} + d_{51} + d_{61} + d_{71} + d_{81}) \\ &= f_3^{(9)} + f_3^{(10)} + f_3^{(8)} + f_3^{(9)} + f_3^{(7)} + f_3^{(9)} + f_3^{(9)} + f_3^{(8)} = 2.200 \end{aligned}$$

(4) “差”类：

$$\begin{aligned} n_{14} &= f_4(d_{11} + d_{21} + d_{31} + d_{41} + d_{51} + d_{61} + d_{71} + d_{81}) \\ &= f_4^{(9)} + f_4^{(10)} + f_4^{(8)} + f_4^{(9)} + f_4^{(7)} + f_4^{(9)} + f_4^{(9)} + f_4^{(8)} = 0 \end{aligned}$$

则得到

$$n_1 = \sum_{i=1}^4 n_{1i} = 7.556 + 6.143 + 2.200 + 0 = 15.909$$

故

$$R_{11} = \left(\frac{n_{11}}{n_1}, \frac{n_{12}}{n_1}, \frac{n_{13}}{n_1}, \frac{n_{14}}{n_1} \right) = (0.475, 0.386, 0.139, 0)$$

依此类推，可分别得到 R_{12} ，从而求出 R_1 ：

$$R_1 = \begin{bmatrix} R_{11} \\ R_{12} \end{bmatrix} = \begin{bmatrix} 0.475 & 0.386 & 0.139 & 0 \\ 0.519 & 0.376 & 0.105 & 0 \end{bmatrix}$$

同理，可以依次求出 R_2 、 R_3 、 R_4 ：

$$R_2 = \begin{bmatrix} R_{21} \\ R_{22} \\ R_{23} \end{bmatrix} = \begin{bmatrix} 0.327 & 0.389 & 0.284 & 0 \\ 0.427 & 0.396 & 0.177 & 0 \\ 0.179 & 0.230 & 0.322 & 0.269 \end{bmatrix}$$

$$R_3 = \begin{bmatrix} R_{31} \\ R_{32} \end{bmatrix} = \begin{bmatrix} 0.282 & 0.362 & 0.356 & 0 \\ 0.358 & 0.397 & 0.245 & 0 \end{bmatrix}$$

$$R_4 = \begin{bmatrix} R_{41} \\ R_{42} \\ R_{43} \end{bmatrix} = \begin{bmatrix} 0.197 & 0.253 & 0.353 & 0.197 \\ 0.261 & 0.335 & 0.404 & 0 \\ 0.242 & 0.311 & 0.391 & 0.056 \end{bmatrix}$$

5. 进行决策

依据公式 $B = A \times R$ ，分别求出 B_1 、 B_2 、 B_3 、 B_4 。

$$B_1 = A_1 \times R_1 = (0.475, 0.386, 0.139, 0),$$

$$B_2 = A_2 \times R_2 = (0.327, 0.368, 0.322, 0.269)$$

$$B_3 = A_3 \times R_3 = (0.358, 0.397, 0.356, 0),$$

$$B_4 = A_4 \times R_4 = (0.242, 0.311, 0.353, 0.197)$$

对 B_1 、 B_2 、 B_3 、 B_4 进行归一化处理，可得

$$B_1 = (0.475, 0.386, 0.139, 0), \quad B_2 = (0.254, 0.287, 0.250, 0.209)$$

$$B_3 = (0.322, 0.358, 0.320, 0), \quad B_4 = (0.219, 0.282, 0.320, 0.179)$$

由 B_1 、 B_2 、 B_3 、 B_4 及 A 可计算 B ：

$$B = A \times R = (0.358, 0.358, 0.250, 0.209)$$

对 B 进行归一化处理：

$$B = (0.305, 0.305, 0.213, 0.177),$$

$$Z = BG^T = 6.299 \in [6, 8]$$

根据求解结果，由评价等级很容易推出该航保备件比较关键，损耗较严重，总体储备效果良好，适合存储。

第 4 章

基于寿命分布和仿真的武器装备备件需求预测方法



针对武器装备备件样本少、数值小、不确定性强的特点，仿真方法为其提供了一种有效的解决途径。仿真方法主要对备件需求产生的过程进行模拟，统计需求量或评估备件保障策略的效能。系统模型中有很多随机变量，而在系统仿真时对这些随机变量的描述只能用有限的样本数据为依据，仿真结果的置信度与样本量有关。

4.1 基于寿命分布的备件需求预测方法

在武器装备使用阶段早期，由于装备使用时间比较短，装备性能与部队训练还没有进入稳定期，备件的消耗主要来自装备中零部件的故障，各种实际统计数据比较少，此时备件需求预测适用于基于内在因素的预测方法。利用零部件的寿命分布参数，依据给定的寿命分布计算模型来预测这一时期的备件需求量，作为初始备件配置的理论依据和主要参考，从而确保武器装备使用阶段早期备件供应保障的合理性。

4.1.1 模型假设

模型建立的假设条件如下：

- (1) 各工作部件相互独立；
- (2) 备件更换瞬间完成；
- (3) 零部件的寿命服从该寿命分布；
- (4) 备件在储备过程中不失效；
- (5) 对于储备周期较长的武器装备，其累计工作时间取折算后的等效时间，即把武器装备储备时间折算为工作时间；
- (6) 备件保障概率 P 设定为 1，即备件需求量等于备件消耗量。

4.1.2 模型建立

寿命分布是可靠性工程研究和应用的基础。常用的寿命分布如指数分布、正态分布、威布尔分布等都由其参数唯一确定。基于寿命分布的预测模型，利用备件不同的故障分布类型计算备件可靠性数据，基于排队论或更新理论求得备件的最小需求量。具体方法是在保证一定备件保障率的前提下使得备件数量最小，或者说在一定的备件费用约束下使得备件保障率最高。运用基于寿命分布函数的备件需求预测方法，关键是确定零部件的故障分布类型及相关参数。武器装备的发展普遍具有同源性和继承性，同源型号的武器装备备件也大多具有通用性和相似性。一般在以前型号装备上使用过的备件，其寿命分布参数可以通过以前型号使用过程中的现场失效样本数据处理——备件寿命分布的拟合检验和分布函数中参数的估计来获取，而少部分新型备件则可以有针对性地通过寿命加速试验来获取相关参数。这为基于寿命分布的备件需求量预测提供了数据来源和方法可行性保障。

本书主要参照 GJB 4355—2002《备件供应规划要求》中给出的服从指数分布、正态分布、威布尔分布的备件需求量计算模型进行预测。

1. 指数寿命件需求预测模型

1) 部件为不可修复件

该模型主要适用于具有恒定故障率的部件。一般来说，正常使用的电子零部件属于指数寿命件，如印制电路板插件、电子部件、电阻、电容、集成电路等。

若某备件的寿命服从指数分布, 则该备件的消耗量可按下式确定:

$$P(j \leq S) = \sum_{j=0}^S \frac{(N\lambda t)^j}{j!} \exp(-n\lambda t) \quad (4-1)$$

式中: S 为备件消耗量; P 为备件保障概率; n 为备件总数; λ 为零部件故障率; t 为保障时效期内工作时间。

据统计, 寿命服从指数分布的部件约占全部部件的 90%。下面给出指数寿命件备件消耗量的计算模型。

当 $n\lambda t > 5$ 时, 可以用正态分布近似计算, 这时备件消耗量的计算式简化为

$$S = n\lambda t + u_p \sqrt{n\lambda t} \quad (4-2)$$

式中: u_p 为正态分布分位数, 可从 GB 4086.1 统计分布数值表正态分布中查出。常用的正态分布分位数见表 4-1。

表 4-1 常用的正态分布分位数

P	0.75	0.9	0.95	0.99	0.995
u_p	0.67	1.28	1.65	2.33	3.09

2) 部件为可修复件

由武器装备持续工作时间 t 和备件保障度要求 $R_p(n)$ 计算首次出现缺备件时间的平均时间:

$$\bar{T} = -\frac{t}{\ln R_p(n)} \quad (4-3)$$

计算维修比:

$$K = \frac{\mu}{\lambda_s} \quad (4-4)$$

式中: λ_s 为该零部件的串联故障率, $\lambda_s = L\lambda$; μ 为备件、零部件修复率 (一个修理工)。

按 K 的数值不同, 分三种情况计算备件储备量 n :

(1) 当 $K \approx 1$ 时 ($K \neq 1$), 有

$$n_1 = \frac{-3 + \sqrt{9 + 8(\lambda_s \bar{T} - 1)}}{2} \quad (4-5)$$

计算出的 n_1 为近似解, 将 n_1 代入下式:

$$n_2 = \frac{\lg[\lambda_s \bar{T}(L-1)^2 + (n_1 + 2)(K-1) + 1]}{\lg K} - 2 \quad (4-6)$$

如果计算出的 n_2 与 n_1 值相差较小 (小于 1), 二者中较大的是要求的备件量 n ; 如果算出的 n_2 与 n_1 值相差较大, 可用 n_2 代替式 (4-6) 中的 n_1 , 进一步求第三近似解 n_3 , 直到 n_i 与 n_{i+1} 之差小于 1, 取较大者作为所要求的 n 。

(2) 当 $K \geq 1.5$ 时, 有

$$n_1 = \frac{\lg[\lambda_s \bar{T}(K-1)^2 + 1]}{\lg K} - 2 \quad (4-7)$$

计算出的 n_1 为近似解, 将 n_1 代入式 (4-6) 以后, 求备件储备量 n 的计算方法与 $K \approx 1$ 时相同。

(3) 当 $K < 0.8$ 时, 有

$$n_1 = \left[\lambda_s \bar{T}(1-K) + \frac{1}{1-K} \right] - 2 \quad (4-8)$$

计算出的 n_1 为近似解, 将 n_1 代入式 (4-6) 以后, 求备件储备量 n 的计算方法与 $K \approx 1$ 时相同。

2. 正态寿命件需求预测模型

正态分布是一种综合性的累积损伤模型。零部件在随机应力作用下逐渐耗损老化而强度下降, 当下降到超过临界值时就产生故障。其有两个特点: 一是所受的应力是多种应力的综合作用, 诸应力是相互独立的, 且无一一是起主导作用的; 二是损伤是累积损伤的结果。备件在环境应力 (高温、潮湿、盐雾、有害气体等) 作用下产生腐蚀、表面破坏、老化等, 如果这些应力中没有一个能起主导作用, 而是共同作用的结果, 没有发生随机干扰引起突然破坏, 则会产生正态型的故障。

适用于零部件的故障模式有机械装置磨损的故障、管路和阀门腐蚀的故障、燃料系统喷油嘴磨损的故障、电气绝缘体老化的故障, 以及工艺参数和性能参数超差的故障等, 如轮胎磨损、变压器、电动绕组绝缘、金属疲劳等。

若已知正态寿命件的均值 E 、标准差 σ 、更换周期 (如果是磨损寿命, t 用工作时数计算; 如果是腐蚀、老化寿命, t 可以用日历时数近似), 则在满足备件保障概率 P 的条件下, 该备件的消耗量为

$$S = \frac{t}{E} + u_p \sqrt{\frac{\sigma^2 t}{E^3}} \quad (4-9)$$

从式 (4-9) 很容易推导出式 (4-2), 只需按正态分布近似计算指数分布的条件, 将 $E = \frac{1}{n\lambda}$, $\sigma^2 = \frac{1}{(n\lambda)^2}$ 代入式 (4-9) 即可。

3. 威布尔寿命件需求预测模型

威布尔故障模型产生的过程已经抽象为链条的断裂过程。在链条中每一环节的强度是不均匀的。若把外力缓慢地加到完整的链条两端,则每一环节上所受的力是相等的。当外力超过最薄弱环节的强度时,该环节就被拉断,于是整根链条被拉断。像这一类零部件的故障时间服从威布尔分布。

威布尔分布是瑞典科学家威布尔提出的,一般比较适用于机电零部件的寿命分布描述,它的导出是基于最弱环原理。很多分布都可以用威布尔分布进行描述,当威布尔形状参数 $\beta < 1$ 时,可以描述伽马分布;当 $\beta = 1$ 时,可以描述指数分布;当 $\beta = 2$ 时,可以描述瑞利分布;当 $\beta = 1.44$ 时,近似于正态分布。在事先不知某零部件寿命服从何种分布时,可以用威布尔分布近似分析处理。

该模型主要适用于装备上的机电件,如继电器、开关、断路器、电容器、电子管、电位计等。

若某零部件寿命分布为三个参数威布尔分布,则有

$$F(t) = 1 - \exp \left\{ - \left[\frac{(t - t_0)}{\alpha} \right]^\beta \right\} \quad (t \geq t_0, \alpha > 0, \beta \geq 0) \quad (4-10)$$

式中: β 为形状参数; t_0 为位置参数; α 为尺度参数。

若有 M 台装备,每台装备上装机 N 个,已知该零部件可修且修复率为 μ ($\mu = 0$ 时,表示该零部件不可修),当装机零部件发生故障时能得到该备件的概率为 P 。设此事件服从正态分布,若单台装备累计工作时间为 t ,则在满足备件保障概率 P 的条件下,该备件的需求量为

$$S = \left\lceil \left\lceil L \cdot (1 - \mu) \cdot \frac{t}{m} + u_p \sqrt{L \cdot \frac{t}{m} \cdot (1 - \mu) \left[\mu + (1 - \mu) \cdot \frac{\sigma^2}{m^2} \right]} \right\rceil \right\rceil \quad (4-11)$$

式中: $L = M \cdot N$; u_p 为正态分布分位数,可从 GB 4086.1 统计分布数值表正态分布中查出,常用的见表 4-1; $\lceil [x] \rceil$ 为不小于 x 的正整数; m 、 σ^2 分别为 $F(t)$ 对应的均值和方差,表达式分别为

$$m = \alpha \Gamma(1 + \frac{1}{\beta}) + t_0, \quad \sigma^2 = \alpha^2 [\Gamma(1 + \frac{2}{\beta}) - (\Gamma(1 + \frac{1}{\beta}))^2]$$

4. 对数寿命件需求预测模型

零部件在交变应力作用下会产生疲劳裂纹,裂纹增加到临界值时会使产品变形甚至断裂。研究证明,裂纹长度(或深度)的增长是按比例递增的。设在相同的应力作用下在相等的时间内,用 $x_1 < x_2 < \dots < x_n$ 表示裂纹的长度,则有

$$\Delta x_i = x_i - x_{i-1} = K_i x_{i-1} \tag{4-12}$$

式中： K_i 是大于 1 的系数。

式(4-12)表示第 i 阶段的裂纹长度增量 Δx_i 同该阶段开始时的裂纹长度 x_i 成比例，故这种故障模型称为比例效应模型。

零部件故障是由各种应力累积损伤引起的，而且损伤是按比例效应累积的，则其故障都属于对数正态分布。构件中的夹渣、裂纹、尖角都是疲劳的敏感因素，但是在一定条件下都会引起累积损伤。如果敏感因素引起的累积效应占主导地位，则产生对数正态分布的故障，如电机绕组绝缘、半导体器件、硅晶体管、弹体结构、金属疲劳等。

5. 预测模型的适用范围

各类分布类型的大致应用范围见表 4-2。

表 4-2 各类分布类型的大致应用范围

分布类型	应用范围
指数分布	对于元件,适用于偶然故障与使用事件无关的情况。常常适用于电子设备和电子元件,多个部件组成的复杂系统,某些软件的故障类型;具有恒定故障率的零部件,经试验并进行定期维修的部件
正态分布	在非寿命情况下,80%以上的问题是整体分布问题。在机械产品和结构工程中用来研究强度和应力分布;磨损型的零件和故障,如齿轮、密封环、轴承、活塞环等故障模式分布;质量过程稳定性时产品尺寸及性能的量度;电阻值、半导体材料性能指标及材料强度;许多作用相当而又相互独立的因素叠加引起的随机变量
对数正态分布	寿命现象事件集中发生在顶部时的不对称情况,且观测值的离散程度很大,如半导体器件、硅晶体管、锗晶体管、风扇叶片、车体结构、金属疲劳等
威布尔分布	适用于有薄弱环节的模型,如机械中的疲劳强度、磨损寿命等,以及滚动轴承、传动齿轮、电动机、发动机等许多机械电气元件等

4.1.3 示例分析

1. 不可修件指数分布件计算举例

某型导弹装备上有一关键组件采用某类型的电子件 12 个，其中 1 个失效则整个组件失效，且为不可修件。已知电子器件的故障率 $\lambda = 4 \times 10^{-4} / \text{h}$ ，经折算该关键组件的累计工作时间约为每年 1000h，求 1 年内保障率不小于 90%的条件下需要多少个备件？

解： $n\lambda t = 12 \times \frac{1}{2500} \times 1000 = 4.8$

通过泊松分布表可查得数据计算备件保障度 U ，如表 4-3 所列，其中 $V = \frac{(n\lambda t)^k}{k!} \exp(-n\lambda t)$ ， $U = \sum_{k=0}^s \left[\frac{(n\lambda t)^k}{k!} \exp(-n\lambda t) \right]$ 。

表 4-3 不同备件量对应的备件保障度

$n\lambda t = 4.5$			$n\lambda t = 5.0$		
备件量 k	V	U	备件量 k	V	U
0	0.0111	0.0111	0	0.0067	0.0067
1	0.0500	0.0611	1	0.0337	0.0404
2	0.1124	0.1736	2	0.0842	0.1247
3	0.1687	0.3422	3	0.1404	0.2650
4	0.1898	0.5314	4	0.1755	0.4405
5	0.1708	0.7022	5	0.1755	0.6160
6	0.1281	0.8303	6	0.1462	0.7622
7	0.0824	0.9127	7	0.1045	0.8666
8	0.0463	0.9590	8	0.0653	0.9319

由表 4-3 可见，当备件数 $s=8$ 时，备件保障度 $U>0.9$ 。

也可利用式（4-2）直接计算结果，见表 4-4。

表 4-4 程序运行结果

备件数量	0	1	2	3	4	5	6	7	8
保障率	0.0082	0.0477	0.1425	0.2942	0.4763	0.6510	0.7908	0.8867	0.9442

同样可得，当备件数 $s=8$ 时，备件保障度 $U>0.9$ 。

2. 可修件指数分布件计算举例

某舰艇编队出海巡航 3 个月，巡航期间不补给。其舰上短波综合通信系统有某部件 12 个同时工作，其单件故障率为 0.005，又随艇修理工修复时，修复率 $\mu=0.07/\text{h}$ ，若要求该艇巡航执行任务期间该部件的备件保障度不小于 90%，试确定备件的储备量 n 。

解： $t = 3 \times 30 \times 24 = 2160\text{h}$ ， $R_p(n) = 0.9$

$$\bar{T} = -\frac{t}{\ln R_p(n)} = \frac{2160}{0.10536} = 20501(\text{h})$$

$$\lambda_s = L\lambda = 12 \times 0.005 = 0.06(1/\text{h})$$

$$K = \frac{\mu}{\lambda_s} = \frac{0.07}{0.06} \approx 1.1667$$

当 $k \approx 1$ 时, 有

$$\begin{aligned} n_1 &= \frac{-3 + \sqrt{9 + 8(\lambda_s \bar{T} - 1)}}{2} \\ &= \frac{-3 + \sqrt{9 + 8 \times (0.06 \times 20501 - 1)}}{2} = 48.1 \\ n_2 &= \frac{\lg[\lambda_s \bar{T} (L-1)^2 + (n_1 + 2)(K-1) + 1]}{\lg K} - 2 \\ &= \frac{\lg[0.06 \times 20501 \times (1.1667 - 1)^2 + (48.1 + 2) \times (1.1667 - 1) + 1]}{\lg 1.1667} - 2 \\ &= 21.4 \\ n_3 &= \frac{\lg[\lambda_s \bar{T} (L-1)^2 + (n_2 + 2)(K-1) + 1]}{\lg K} - 2 \\ &= \frac{\lg[0.06 \times 20501 \times (1.1667 - 1)^2 + (21.4 + 2) \times (1.1667 - 1) + 1]}{\lg 1.1667} - 2 \\ &= 20.7 \end{aligned}$$

因为 n_3 与 n_2 之差小于 1, 不必再进行计算, 二者中较大的即要求的备件量。为保险起见取备件储备量 n 为 22 件。

3. 正态寿命件计算举例

已知某型武器装备中某正态寿命件的平均寿命 $E = 10^3 \text{ h}$, 方差 $\sigma^2 = 200^2 \text{ h}^2$, 更换周期 $t = 2 \times 10^4 \text{ h}$, 求备件保障率 $U = 0.95$ 时的备件需求量。

解: 由表 4-1 可知, $U = 0.95$ 对应的 $u_U = 1.65$, 可得

$$S = \frac{t}{E} + u_U \sqrt{\frac{\sigma^2 t}{E^3}} = \frac{2 \times 10^4}{10^3} + 1.65 \times \sqrt{\frac{200^2 \times 2 \times 10^4}{10^9}} \approx 21.5$$

通过取整得该部件需要 22 个备件。

4. 威布尔寿命件计算举例

已知某型导弹舵系统中有某电动舵机部件 $M=5$ 个, 该部件中某威布尔寿命件器件的机用数 $N=3$ 个, 已知参数 $\alpha=1000\text{h}$, $\beta=2$, $t_0=0$, 单部电动舵机的工作时间 $T=5000\text{h}$, 备件的修复率 $\mu=0.90$, 备件的保障度 $U=0.95$, 并设此事件 (“备件故障时能得到满足”) 服从正态分布, 共需备件多少件?

解: 由表 4-1 可知, $U=0.95$ 对应的 $u_p=1.65$, 已知 $\beta=2$, 查 Γ 函数值得

$$\Gamma(1+\frac{1}{\beta})=\Gamma(1+0.5)=0.8862$$

$$\Gamma(1+\frac{2}{\beta})=\Gamma(1+1)=1.000$$

$$L=M \cdot N=15$$

$$m=\alpha\Gamma(1+\frac{1}{\beta})+t_0=1000 \times 0.8862+0=886.2$$

$$\sigma^2=\alpha^2[\Gamma(1+\frac{2}{\beta})-(\Gamma(1+\frac{1}{\beta}))^2]=1000^2 \times [1-0.8862^2]=1000^2 \times 0.2146$$

代入式 (4-11) 可得

$$S=\left[\left[15 \times (1-0.9) \times \frac{5000}{886.2}+1.65 \times \sqrt{15 \times (1-0.9) \times [0.9+(1-0.9) \times \frac{1000^2 \times 0.2146}{886.2^2}]} \right]\right]$$

$$=15$$

因此, 该部件需要备件 15 件。

基于寿命分布函数的武器装备备件需求预测方法科学性和逻辑性强, 对于故障分布类型可以确定的备件, 预测效果比较好, 但需要掌握每种备件的可靠性及维修性参数和指标, 如备件的寿命分布类型、可靠度、维修度等参数, 以及平均寿命、平均故障间隔时间和平均修复时间等。这些参数和指标往往不易得到, 因而难以进行精确计算。

4.2 基于蒙特卡罗仿真的备件需求预测方法

传统的备件需求预测方法是在已有备件需求历史数据的基础上, 对备件历史数据进行统计和分析, 从而对下一步的备件需求量进行预测。但武器装备研制阶

段和使用阶段早期,由于缺少备件消耗的历史数据,无法用传统的方法进行预测。基于仿真方法的备件需求预测是在充分考虑备件需求影响因素的基础上,运用系统仿真技术对零部件的使用寿命过程进行仿真模拟,从而能够在最贴近实际的情况下得出备件需求预测值。

4.2.1 模型的假设

在实际中,备件需求量除受自身可靠性影响外,还受到很多其他因素影响。通常包括部件自身的故障率、部件的易损程度、装备的使用环境、装备的使用强度和装备的增减。此外,还存在一些不确定、未可知的偶然因素的影响,以及武器装备的发展变化的影响等。因此,备件需求量的分布往往不是一个线性分布,部件的故障率曲线也不一定能找到合适的解析式来描述,要建立一个能够较全面描述它的数学计算模型比较困难。因此,采用仿真方法对备件需求量进行预测是可行的。

在运用仿真方法进行备件需求预测前,需提出下列假设:

(1) 一个部件的实际使用寿命以它装配到装备上的时刻开始,到它发生故障、不能实现预定功能的时刻为止。

(2) 多个部件失效为独立事件,在同一时刻不会发生多个事件。

(3) 部件发生故障后,直接用备件替换,不考虑当时修复继续使用的可能性。

(4) 部件一经发生故障,就认为该部件已被换下,即更换备件花费时间为零。

(5) 部件的故障率曲线为 $\lambda_s(t)$, 且在不同的故障期只考虑一种主要的故障率。

(6) 部件的易损程度导致的消耗率曲线为 $f_c(t)$ 。

(7) 装备的使用环境导致的消耗率曲线为 $f_a(t)$ 。

(8) 装备的使用强度导致的消耗率曲线为 $f_e(t)$ 。

(9) 装备增减导致的备件消耗率为零。装备增减事件的发生是明确的,其变化对备件消耗的影响也相对比较明确,所以,在一般情况下不予考虑。

(10) 不确定、未可知的因素导致的备件消耗率为零。由于这些因素发生概率很小,具有突发性,在一般情况下不予考虑。

4.2.2 参数确定

确定各部件寿命的随机变量分布类型及其参数的方法可由该部件或同类部件寿命历史数据经统计方法得到,部分没有历史数据的部件可参照零件出厂说

明。多重故障、人为差错、环境因素等因素对部件影响程度系数,可由因素分析法、层次分析法、粗糙集法来确定。

4.2.3 蒙特卡罗方法

蒙特卡罗方法是一种计算机模拟方法,又称为随机模拟法或统计试验法,是一种根据统计抽样理论近似求解数学问题或物理问题的方法,能够比较逼真地描述具有随机性质的事物的特点及物理实验过程,可以部分代替物理实验,甚至可以得到物理实验难以得到的结果,现已广泛应用于工程领域。

具体来说该方法是针对待求问题,根据物理现象本身的统计规律,或人为构造一合适的依赖随机变量的概率模型,使某些随机变量的统计量为待求问题的解,进行大统计量 $N \rightarrow \infty$ 。其思路是:首先建立一个与所求解有关的概率模型或者随机过程,使所求的解就是所建立模型的概率分布或数学期望;然后对这个模型进行随机抽样观察,即产生随机变量;最后进行统计模拟,求得近似分布的估计值,以作为近似解,解的精度可用估计值的标准误差来表示。其方法简单、便于编制计算机程序,能够保证依概率收敛,而且精度较高。随着计算机技术的发展和普及,这种技术越来越获得广泛应用,已经可以求解由任意分布类型的部件组成的复杂系统的最优备件需求量。

用该方法解决实际问题,可以直接从该问题本身出发,而不从方程或数学表达式出发,它有直观、形象的特点。此外,在计算机上进行蒙特卡罗仿真计算时,程序结构简单,分块性强,易于实现。蒙特卡罗方法的最大优点是:其计算的收敛速度和误差的大小与问题的复杂程度或相空间的维数无关,可以很好地避免“维数魔咒”限制,这一性质使得它在科学和工程的许多领域获得了广泛应用。

4.2.4 仿真模型建立

假设系统中共使用 m 个某类部件,且各个部件已经工作过的时间长短不完全相同。在仿真过程中,将一个时间段按一定的时间步长分为若干个小时的时间片段,使用故障率曲线和影响因素导致的消耗率曲线逐个模拟部件在每个小时时间片段中的故障是否发生事件,再进行数量统计。在使用仿真方法预测备件需求量时,至少对每个位置的部件均仿真一次,共仿真 m 次。

在仿真模型中,首先输入各部件的故障率及其他因素导致的备件消耗率数据,并初始化仿真变量。仿真的基本事件简化为故障,仿真推进点为故障点,利

用蒙特卡罗抽样技术获取故障件发生时间，采用事件调度法的仿真策略对第 $k(1 \leq k \leq m)$ 个部件建立仿真模型如图 4-1 所示。

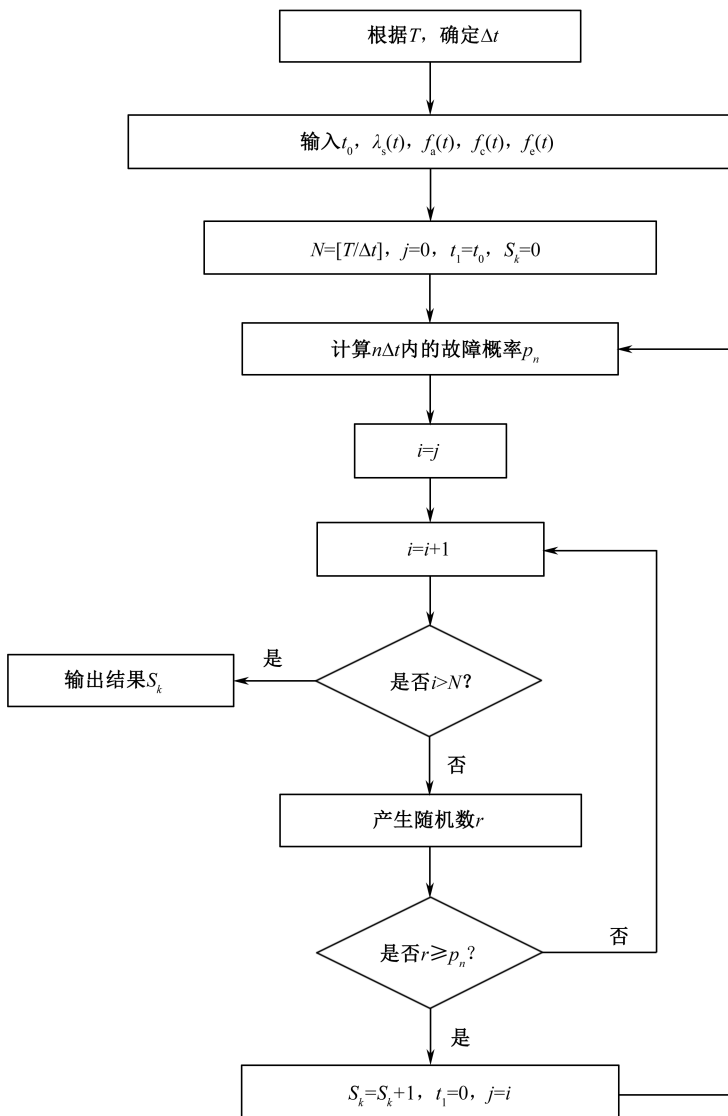


图 4-1 备件消耗预测仿真模型

图 4-1 中： T 为未来时间段； j 为已仿真次数； n 为 Δt 的数目， $n = N - j$ 且有 $0 \leq n \leq N$ ； Δt 为仿真步长； t_0 为仿真前该部件已经工作过的时间； S_k 为一次

仿真结果： p_n 为部件在未来一段时间内发生故障的概率，且有

$$p_n = 1 - (1 - p_s)(1 - p_c)(1 - p_a)(1 - p_e) \quad (4-13)$$

式中： $p_s = \int_{t_1}^{t_1+n\Delta t} \lambda_s(t)dt$ ； $p_c = \int_{t_1}^{t_1+n\Delta t} f_c(t)dt$ ； $p_a = \int_{t_1}^{t_1+n\Delta t} f_a(t)dt$ ；

$$p_e = \int_{t_1}^{t_1+n\Delta t} f_e(t)dt。$$

第 k 个部件的随机数 r 是服从某一概率分布的随机变量，生成随机数的方法有逆变法、函数变换法、卷积法、取舍法和组合法等，这里采用逆变法。以服从指数分布的部件为例， μ 是 (0,1) 区间内均匀分布的随机变量，且有 $\lambda = \lambda_s(t) + f_c(t) + f_a(t) + f_e(t)$ 。

在对每个部件都进行了仿真共 m 次之后，可得到未来时间段 T 内 m 个部件对应的备件消耗量为

$$S = \sum_{k=1}^m S_k \quad (4-14)$$

以每组仿真 m 次的数目重复进行 D 组仿真，假设第 d 组仿真得到的备件需求量为 $S_d (d=1,2,\dots,D)$ ，则该类备件需求量预测值为

$$S_d = \bar{S} \pm 3\sigma \quad (4-15)$$

式中

$$\bar{S} = \frac{1}{D} \sum_{d=1}^D S_d; \sigma = \sqrt{\frac{1}{D-1} \sum_{d=1}^D (\bar{S} - S_d)^2}$$

对备件需求量预测值 S_d 进行取整处理后得到 S'_d ：

$$S'_d = \begin{cases} S_d & (S_d = [S_d]) \\ [S_d] + 1 & (S_d > [S_d]) \end{cases} \quad (4-16)$$

4.2.5 示例分析

现有 3 台某型导弹测试设备，每台装备上装有 3 个保险管，把该部件的有用寿命期划分为 4 个阶段。这 9 个部件已经工作时间分别为 0.05, 0.31, 0.27, 0.42, 0.02, 0.23, 0.19, 0.2, 0.14，且在前 5 个时间段内的备件消耗量分别为 2, 10, 3, 7, 6，预测第 6 个时间段内的该备件的需求量。

解：部件的故障率曲线为浴盆曲线 $\lambda(t)$ ，有

$$\lambda_s(t) = \lambda_1(t) + \lambda_2(t) + \lambda_3(t)$$

式中

$$\lambda_1(t) = 0.3375te^{-0.9t} (0 \leq t \leq 1)$$

$$\lambda_2(t) = 0.3375e^{-0.9} (1 \leq t \leq 2)$$

$$\lambda_3(t) = 0.3375e^{-0.9} t^2 / 4 (2 < t \leq 4)$$

同时, 由易损程度、使用环境和使用强度造成部件故障的消耗率曲线分别为

$$f_a(t) = 1 - e^{-0.1t} (t \geq 0)$$

$$f_c(t) = 1/100 (t \geq 0)$$

$$f_c(t) = 0.2e^{-0.1t} (t \geq 0)$$

根据式 (4-13) 计算故障概率 p_n , 运行 Matlab 程序进行仿真。在仿真过程中, 所取的仿真时间步长 $\Delta t = 0.05$ 。根据式 (4-14) ~ 式 (4-16) 计算得, 该部件备件的需求量为 16 个。

4.3 基于马尔可夫与蒙特卡罗仿真的备件需求预测方法

间断消耗件也称慢速流动件, 是指该件一般不发生故障, 在历史消耗数据中往往表现出大量的零值。在武器装备保障过程中, 发现有一些备件属于使用频率低、间隔期长且消耗规律不确定的低消耗备件, 该类备件一般具有可用性要求高、专用性强、单价高、功能关键等特点。通过对此类备件历史消耗数据的观察, 发现它们通常是存在大量零值的非负整数序列, 而这些非负整数都是随机出现的, 且数值不大。这种间断性的低消耗特征使得其预测非常困难, 用一般的连续性预测方法进行预测精度很低。又由于样本量极少, 靠传统的现场统计数据来确定这类备件的寿命分布和分布参数几乎不可能, 可见基于可靠性的预测方法也不适用。鉴于以上分析, 将马尔可夫链与蒙特卡罗仿真相结合, 构造一种间断性低消耗备件需求预测模型。

4.3.1 马尔可夫预测模型

1. 马尔可夫链预测思想

马尔可夫链预测思想由马尔可夫创立, 是一种分析随机过程的预测方法。具体来说, 马尔可夫链预测模型的主要研究对象是一个运行系统的状态和状态转移。该系统是未来状态仅与当前状态有关, 而与历史状态无关的随机过程, 即马尔可夫性。其原理是根据系统当前状态及变化趋势, 通过马尔可夫链理论计算状态转移概率预测系统未来达到的某些状态的概率。1 个 n 阶马尔可夫链由 n 个状

态集合 (E_1, E_2, \dots, E_n) 和一组转移概率 $p_{ij}(i, j = 0, 1, 2, \dots, n)$ 确定, 该系统在任意一个时刻只能处于一种状态。若在时刻 k , 过程处在状态 E_i , 则在时刻 $k+1$, 它将以概率 p_{ij} 处于状态 E_j , 转移概率 p_{ij} 也就反映了系统随机因素影响的程度。

2. 马尔可夫过程

马尔可夫链预测模型对未来状态的预测, 不需要寻求系统复杂因素之间的相互规律, 只需考虑系统本身历史状态的演变特点, 是一种动态随机数学模型。

马尔可夫过程是马尔可夫理论的精髓, 是实际应用中最重要随机过程, 在复杂随机过程建模中具有特殊的意义和作用。动态系统在每个时期所处的状态是随机的, 从一个状态转移到下一个状态按照一定的概率进行, 并且下一个时期的状态和转移概率与以前各时期的状态无关, 这种性质称为无后效性或马尔可夫性。具有马尔可夫性的时间和状态均为离散的随机转移过程, 可用马尔可夫过程描述。

马尔可夫过程是在已知当前状态的情况下, 过程的未来状态与其过去状态无关的随机过程。马尔可夫链预测理论通过使用状态转移矩阵来确定状态的变化趋势, 从而达到对未来状态进行预测的目的。马尔可夫链预测的对象是一个随机变化的动态系统, 预测要求的状态数量较少, 例如, 某量“增加”或“减少”, 企业经营状况“好”或“坏”, 对于长期预测和随机波动较大的数据序列的预测效果较好, 对解决无后效性的预测有独到之处。马尔可夫链预测理论的研究成果较多, 在教育、经济学、金融投资、生物学、农业、水资源科学中得到广泛应用。

3. 马尔可夫链预测模型

设 X_t 是 t 时刻事物所处的状态, 为便于研究, 首先将连续变化的时间进行“离散化”处理; 同时将事物所属的状态分成若干等级, 分别与状态空间对应。马尔可夫模型建立的基础是“无后效性”和“平稳性”, 它决定了此模型的适用范围, 即应用马尔可夫预测法必须满足状态转移概率矩阵具有一定的稳定性。无后效性是指事物本阶段的状态只与前一个阶段的状态有关, 而与以前其他任何阶段的状态无关, 即对于任意的正整数 n , 以及 $i, j, i_k(k=0, 1, 2, \dots, n-1)$, 有

$$P\{X_{n+1}=j|X_n=i, X_{n-1}=i_{n-1}, \dots, X_0=i_0\}=P\{X_{n+1}=j|X_n=i\}$$

如果上式条件概率与 n 无关, 即

$$P\{X_{n+1}=j|X_n=i\}=\dots=P\{X=j|X_0=i\}=P_{ij}^{(1)}(n=1, 2, \dots)$$

则称马尔可夫链具有“平稳性”。 $P_{ij}^{(1)}$ 表示由状态 i 经过一步转移到状态 j 的概率, 或记为 P_{ij} 。 P_{ij} 具有下列性质:

$$\begin{cases} P_{ij} \geq 0 & (i, j \in E) \\ \sum_{j \in E} P_{ij} = 1 & (i, j \in E) \end{cases} \quad (4-17)$$

以 P_{ij} 为元素的矩阵 $\mathbf{P} = P_{ij}$ 称为状态转移概率矩阵，其形式为

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix} \quad (4-18)$$

式中：矩阵 \mathbf{P} 中元素 (P_{ij}) 仍具有性质 (4-17)。

由于马尔可夫过程满足“无后效性”和“平稳性”，因此

$$\mathbf{P}^{(k)} = \mathbf{P}^{(k-1)} \cdot \mathbf{P} = \mathbf{P}^{(k-2)} \cdot \mathbf{P}^2 = \cdots = \mathbf{P}^k \quad (4-19)$$

式中： $\mathbf{P}^{(k)}$ 为事物的状态经过 k 步转移的转移概率矩阵。

4.3.2 备件消耗的马尔可夫性分析

低消耗备件是指其每一年（或季度）的消耗量 $X \leq M$ （ M 为一个较小的整数）。大多数低消耗备件的消耗量 X 是一个随机变量，它在一定的范围内呈现出随机的波动性。因素分析法及趋势外推法均不能有效地预测其消耗量。因此，可以假设这一类备件的消耗量与消耗历史无关，只与当前状态有关，即消耗量具有马尔可夫性。

马尔可夫性也称无后效性，即若 $t_1, t_2, \cdots, t_{n-2}$ 表示过去时刻， t_{n-1} 表示现在时刻，低消耗备件消耗过程 $\{X(t), t \in [0, +\infty]\}$ 在 $t_1, t_2, \cdots, t_{n-1}$ 的取值分别为 $x_1, x_2, \cdots, x_{n-1}$ ，则将来时刻 t_n 的状态 $X(t_n)$ 的统计特性取决于现在时刻 t_{n-1} 的状态，与过去时刻无关。

$X = \{x_1, x_2, \cdots, x_n\}$ ，令 $X_{\max} = \max\{x_i | i = 1, 2, \cdots, n\} = M$ ，并假定 $X > M$ 为小概率事件，可以忽略，则季度消耗量有 $M+1$ 种状态： $\{0, 1, 2, \cdots, M\}$ 。用向量 \mathbf{C} 来表示当前季度各种状态出现的概率，记为 $\mathbf{C} = (c_0, c_1, \cdots, c_M)$ 。若当前季度的消耗量 $X_n = i$ ，则 $c_i = 1$ ， $c_j = 0$ ， $i, j \in \{0, 1, 2, \cdots, M\}$ 且 $i \neq j$ ，于是可以通过求转移概率矩阵来推导未来几个季度各种消耗量状态出现的概率。对未来四个季度备件消耗量各种状态的概率求和可以得到未来一年备件消耗量的概率分布。

4.3.3 基于马尔可夫与蒙特卡罗仿真预测的模型建立

1. 模型的建立

一般情况下,一步转移概率的理论分布是未知的,但当样本足够大时,可近似地用状态相互转移的频率来描述。设 p_{ij} 为状态 i 转移到状态 j 的样本个数,则有

$$p_{ij} = P\{i \rightarrow j\} = \frac{A_{ij}}{A_i} \quad (i, j \in \{1, 2, \dots, M\}) \quad (4-20)$$

式中: A_i 为处于状态 i 的样本个数; A_{ij} 为样本中状态 i 一步转移到状态 j 的个数。

从而可以由样本序列 X 获得一步转移概率矩阵为

$$\mathbf{P}_{(1)} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0M} \\ p_{10} & p_{11} & \cdots & p_{1M} \\ \vdots & \vdots & & \vdots \\ p_{M0} & p_{M1} & \cdots & p_{MM} \end{bmatrix} \quad (4-21)$$

由随机过程理论可知以下结论:

$$p_{ij}(n\Delta) = p_{ij}[(n-1)\Delta] \cdot p_{ij}(\Delta) \quad (4-22)$$

或

$$p_{ij}^n = p_{ij}^{(n-1)} \cdot p_{ij} \quad (4-23)$$

$$p_{ij}^n = (p_{ij})^n \quad (4-24)$$

系统的 n 步转移矩阵可以由 $n-1$ 步转移矩阵乘上一步转移矩阵求得,也可由一步转移矩阵的 n 次方求得。因此

$$\mathbf{P}_{(n)} = \mathbf{P}_{(1)}^n \quad (4-25)$$

用 y_k 表示未来 k 个季度各种消耗量状态概率的预测值,则可以建立预测模型如下:

$$\mathbf{y}_k = \mathbf{C} \cdot \mathbf{P}_{(1)}^k = \{a_{k0}, a_{k1}, \dots, a_{kM}\} \quad (4-26)$$

假定当前季度序列号为 l , 则未来四个季度的总消耗量 t 的预测值为

$$t = \sum_{k=1}^4 y_{(l+k)} \quad (4-27)$$

由式 (4-27), 应用蒙特卡罗仿真可得年消耗量的概率分布为

$$p_y = \{p_y(i) | i = 0, 1, \dots, 4M\} \quad (4-28)$$

若用备件满足率来衡量备件的供应水平，则备件满足率可表示为

$$R = \sum_{i=0}^S p_y(i) \quad (4-29)$$

式中： S 为备件需求量。根据设定的备件满足率 R 就可求得备件需求量 S 。

2. 预测步骤

针对间断性的低消耗武器装备备件，建立基于马尔可夫链与蒙特卡罗仿真相结合的预测模型，其基本步骤如下（详细流程见图 4-2）：

（1）提出间断性低消耗备件需求预测问题，构思建立基于马尔可夫链与蒙特卡罗仿真的预测模型。

（2）对间断性低消耗备件的需求量进行马尔可夫性分析。

（3）进行马尔可夫链预测。确定消耗量的状态数，求解当前季度各种状态出现的概率，通过求转移概率矩阵来推导未来四个季度各种消耗量状态出现的概率，从而确定未来一年备件需求量的概率分布。

（4）进行蒙特卡罗仿真求得年消耗量的概率分布。

（5）预测备件下一年的需求量。

3. 仿真步骤

在建立仿真模型的基础上，以 Matlab 为平台，给出了利用计算机实现蒙特卡罗仿真的算法步骤。具体步骤如下：

（1）输入间断性低消耗备件消耗的历史数据组成的时间序 $X = \{x_1, x_2, \dots, x_n\}$ ，确定消耗量的最大值 M ，即 $X_{\max} = M$ 。设定初始值，令仿真次数的初始值 $m = 0$ ，年消耗个数为 i 的概率 $p(i) = 0$ ， $i = (0, 1, \dots, 4M)$ ，同时记录最近季度备件的消耗数 h ，并设向量 $C = (0, 0, \dots, 1, \dots, 0)$ ，其中‘1’位于第 $h+1$ 列。

（2）输入一步转移概率矩阵 $p_{(1)}$ ，分别计算出二、三、四步转移矩阵 $p_{(2)}$ 、 $p_{(3)}$ 、 $p_{(4)}$ ，根据向量 C 的值，求出 a_{kn} 的值。

（3）确定初始值，令 $m = m + 1$ ，利用 Matlab 产生随机数命令 rand(4,1) 产生四个在 (0,1) 上均匀分布的随机数 r_1, r_2, r_3, r_4 。

（4）确定未来第一个季度的备件消耗（ $k=1$ ）：

若 $0 < r_1 < a_{10}$ ，则 $x_1 = 0$ ；

若 $a_{10} < r_1 \leq a_{10} + a_{11}$ ，则 $x_1 = 1$ ；

若 $a_{10} + a_{11} < r_1 \leq a_{10} + a_{11} + a_{12}$ ，则 $x_1 = 2$ ；

若 $a_{10} + a_{11} + a_{12} < r_1 \leq a_{10} + a_{11} + a_{12} + a_{13}$ ，则 $x_1 = 3$ 。

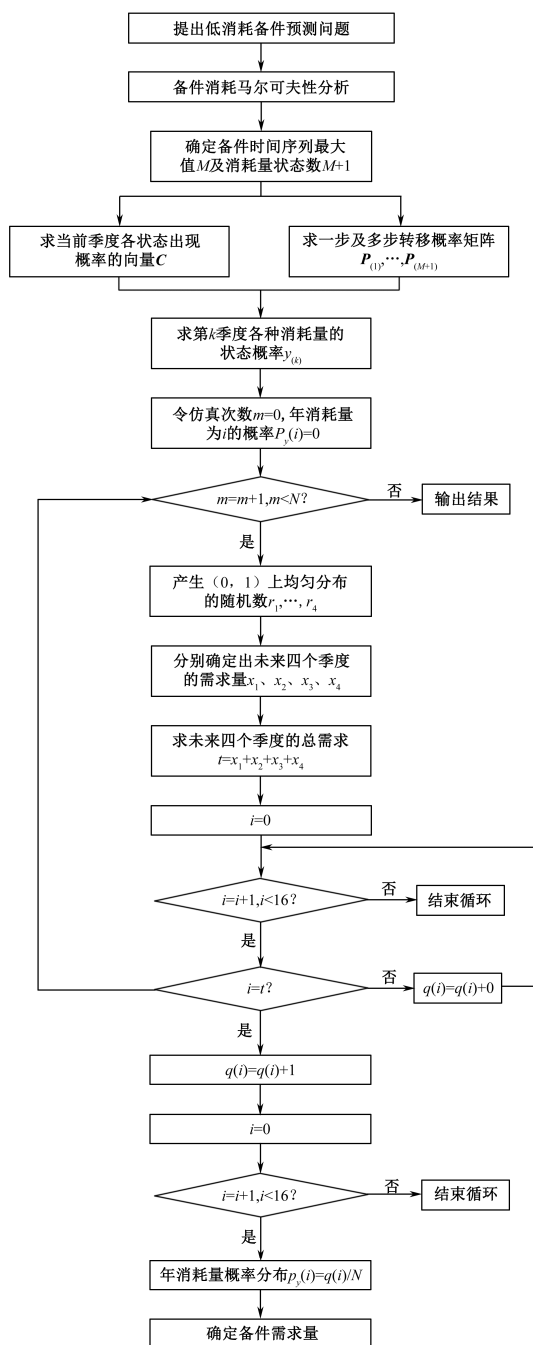


图 4-2 基于马尔可夫和蒙特卡罗仿真的备件需求量预测流程

(5) 同步骤 (4) 可预测未来第二、三、四季度的备件需求量 x_2 、 x_3 、 x_4 。

(6) $p(i) = p(i) + f(t=i)$ 。其中： $i=0,1,2,\dots,4M$ ； t 为四个季度的总需求量， $t = x_1 + x_2 + x_3 + x_4$ ； $f(t=i)$ 为条件函数，若 $t=i$ ，则 $f(t=i)=1$ ；反之，则等于 0。

(7) 当 $m < N$ 时 (N 为设定的仿真次数)，仿真尚未结束时，转至步骤 (2)，否则仿真结束，进入下一步骤。

(8) 仿真结束，输出计算结果 $p_y(i) = p(i)/N$ 。

4.3.4 示例分析

某型自锁螺母是某型导弹装备的零部件，2003—2007 年度中的 20 个季度中，该部件备件的消耗量 $X = \{0, 3, 0, 2, 1, 2, 1, 2, 0, 3, 1, 0, 0, 1, 2, 0, 3, 1, 0, 1\}$ ，由 X 可知 $M = \max_{1 \leq k \leq 20} \{X_k\} = 3$ ，对时间序列进行统计，用频率近似地表示一步转移概率，可得 $P_{ij}(i, j=0, 1, 2, 3)$ 。在序列 X 中，状态为 0 的样本个数 $A_0 = 7$ ，由状态 0 一步转移到状态 0 的样本个数 $A_{00} = 1$ ，则 $p_{00} = P\{0 \rightarrow 0\} \approx \frac{A_{00}}{A_0} = \frac{1}{7} = 0.1429$ 。从而得到转移矩阵 $P_{(1)}$ ：

$$P_{(1)} = \begin{bmatrix} 0.1429 & 0.2857 & 0.1429 & 0.4286 \\ 0.3333 & 0.0000 & 0.5000 & 0.0000 \\ 0.5000 & 0.5000 & 0.0000 & 0.0000 \\ 0.3333 & 0.6667 & 0.0000 & 0.0000 \end{bmatrix}$$

依式 (4-25) 可得到矩阵 $P_{(2)}$ 、 $P_{(3)}$ 、 $P_{(4)}$ ：

$$P_{(2)} = \begin{bmatrix} 0.3299 & 0.3980 & 0.1633 & 0.0612 \\ 0.2976 & 0.3452 & 0.0476 & 0.1429 \\ 0.2381 & 0.1429 & 0.3215 & 0.2143 \\ 0.2698 & 0.0952 & 0.3810 & 0.1429 \end{bmatrix}$$

$$P_{(3)} = \begin{bmatrix} 0.2819 & 0.2167 & 0.2462 & 0.1414 \\ 0.2290 & 0.2041 & 0.2151 & 0.1276 \\ 0.3138 & 0.3716 & 0.1055 & 0.1021 \\ 0.3084 & 0.3628 & 0.0862 & 0.1157 \end{bmatrix}$$

$$P_{(4)} = \begin{bmatrix} 0.2827 & 0.2979 & 0.1486 & 0.1208 \\ 0.2508 & 0.2581 & 0.1348 & 0.0982 \\ 0.2554 & 0.2104 & 0.2307 & 0.1345 \\ 0.2466 & 0.2083 & 0.2255 & 0.1322 \end{bmatrix}$$

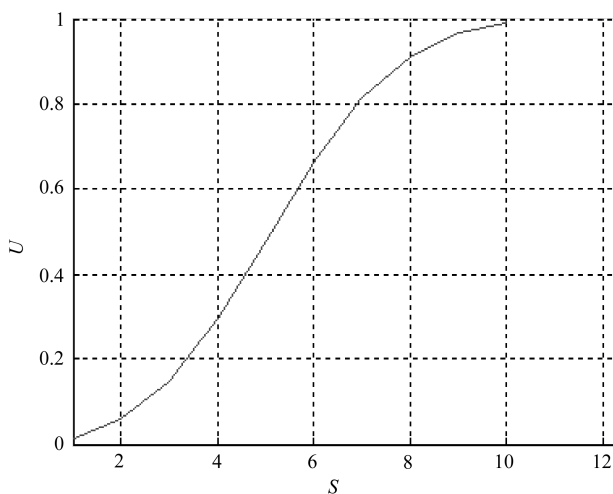
根据消耗序列, 当前月该备件消耗量为 1, 各种状态出现的概率 $C = (0, 1, 0, 0)$, 由式 (4-26) 可分别计算未来 k 季度各种消耗量状态概率的预测值 y_k 。

根据上面给出的仿真步骤在计算机上仿真 10^5 次可得下一年该备件消耗量为 i 的概率 $p_y(i)$, 见表 4-5。

表 4-5 概率对应

i	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
$p_y(i)$	0.014	0.0434	0.0879	0.1499	0.1830	0.1835	0.1515
i	$i=7$	$i=8$	$i=9$	$i=10$	$i=11$	$i=12$	—
$p_y(i)$	0.0978	0.0550	0.0238	0.0076	0.0026	0.0000	—

由此结果可得到该备件消耗个数 S 与该备件满足率 R 的关系如图 4-3 所示。从图可以看出, 当该备件数量为 11 个时, 备件的满足率为 100%, 即备件的消耗量等于需求量。

图 4-3 自锁螺母备件消耗个数 S 与满足率 R 的关系曲线图

4.4 基于 Logistic 回归和马尔可夫过程的备件需求预测仿真模型

备件需求与武器装备可靠性、维修方式、使用环境等因素有关。统计表明, 由装备自身可靠性产生的备件需求通常是平稳需求, 需求序列具有自相关性; 使

用任务、使用环境等因素产生的备件需求是非平稳的，引起非平稳需求的因素在统计学中也称为解释变量。间断性备件需求是两类因素的综合体现，间断性备件需求预测的目的就是要找到备件需求量与解释变量之间的关系及备件需求序列的自相关特性。因此，本书提出先从解释变量序列和自相关序列两个方面估计提前期备件需求发生概率，再通过仿真进一步预测备件的需求数量，从而达到提高预测准确性的目的。

4.4.1 模型建立

备件需求的数学描述如下：设 t 阶段实际发生的需求为 $x(t)$ ($t=1,2,\cdots,n$)，其中有多个阶段的需求为零。当需求为非零时，其数值的大小又具有很大的随机性。预测的目标是估计提前期 L 内的总需求，称为提前期需求量，表示为

$$Y_{\text{LTD}} = \sum_{t=n+1}^{n+L} x_t \quad (4-30)$$

首先将需求时间序列 $\{x_1, x_2, \cdots, x_n\}$ 转换成需求是否发生的时间序列 $\{y_i | i=1,2,\cdots,n\}$ ， $y_i=0$ 表示没有需求， $y_i=1$ 表示发生非零需求；然后区分非零需求发生的原因，需求由自相关原因引起的划入“自相关序列”，需求由解释变量原因引起的划入“解释变量序列”。

采用一阶自相关系数 r 来判断序列 $\{y_1, y_2, \cdots, y_n\}$ 的自相关性，自相关函数定义为

$$r = \frac{1}{n-1} \sum_{i=1}^{n-1} y_i y_{i+1} - \left(\frac{1}{n-1} \sum_{i=1}^{n-1} y_i \right) \cdot \left(\frac{1}{n-1} \sum_{i=1}^{n-1} y_{i+1} \right) \quad (4-31)$$

式中： $r \in [-1/4, 1/4]$ 。当 $r = -1/4$ 时，序列 $\{y_1, y_2, \cdots, y_n\}$ 一阶负自相关；当 $r = 1/4$ 时，序列一阶正自相关；当 $r = 0$ 时，序列不相关。

对二分类因变量（因变量只取两个值）进行回归分析时经常使用的统计分析方法是 Logistic 回归预测，而自相关判别常用的方法是马尔可夫过程预测。因此，预测思想是：首先将备件需求时间序列划分为解释变量序列和自相关序列；其次对解释变量序列用 Logistic 回归模型估计提前期备件非零需求发生概率，对自相关序列用马尔可夫过程预测提前期备件非零需求发生概率；然后对上述的两个预测结果进行整合，便可以得到最终的提前期非零需求发生概率；最后在确定提前期非零需求发生概率的基础上，采用 Bootstrap 方法得到提前期备件需求预测结果。模型建立流程如图 4-4 所示。

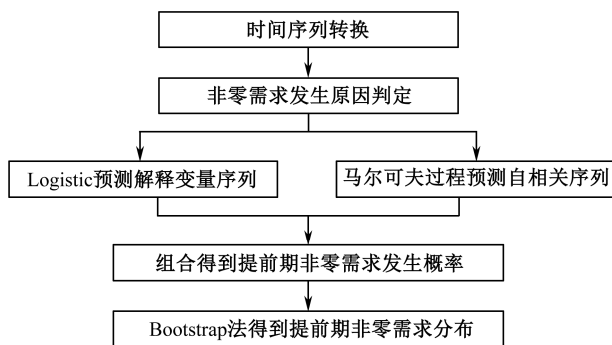


图 4-4 模型建立流程

4.4.2 需求发生概率预测

1. 需求发生原因判定

首先需要建立一种机制来判别非零需求是由需求自相关性引起的，还是由解释变量引起的；然后将由解释变量影响导致的需求和由自相关引起的需求，分别采用相应的方法进行预测。

算法 1 及步骤如下：

- (1) 设定 \bar{r} 的初始值，令 $t=1$ 。
- (2) 当 $t \leq n$ 时，转到步骤 (3)；否则转到步骤 (5)。
- (3) 如果 $y_t = \hat{y}_t = 1$ ，则令 $S_t = \{y_1, y_2, \dots, y_{t-1}, y_t\}$ ， $S'_t = \{y_1, y_2, \dots, y_{t-1}, 0\}$ ，计算 S_t 的相关系数 r_t 和 S'_t 的相关系数 r' ；如果 $|r_t| < \bar{r}$ 或 $|r_t| < |r'|$ ，令 $y_t = 0$ 。
- (4) 令 $\bar{y}_t = y_t$ ， $t = t + 1$ ，转到步骤 (2)。
- (5) 结束。

该算法中，步骤 (1) 中的参数 \bar{r} 可以灵活设定，其目的是平衡解释变量和自相关性对备件非零需求发生的影响。当 $t \in \{1, 2, \dots, n\}$ 时，如果在大部分情况下 $y_t = \hat{y}_t = 1$ ，可以设定较大的 \bar{r} 值（如 $\bar{r} = 0.90$ ）；在半数情况下 $y_t = \hat{y}_t = 1$ ，可以设定中等的 \bar{r} 值（如 $\bar{r} = 0.05$ ）；在很少情况下 $y_t = \hat{y}_t = 1$ ，则设定较小的 \bar{r} 值（如 $\bar{r} = 0.01$ ）。步骤 (3) 是对 $\{\hat{y}_t | t = 1, 2, \dots, n\}$ 序列中非零需求的归属进行判定。如自相关性很强，强于预先设定的界限 \bar{r} ，且将该非零需求变成零以后自相关性变弱，则该非零需求判为自相关序列，否则判为解释变量序列。步骤 (4) 通过将由解释变量引起的非零需求更换为零需求，便得到需求发生的自相关需求时间

序列 $\{\bar{y}_t | t=1, 2, \dots, n\}$ 。

2. 马尔可夫自相关预测

由于 $\{\bar{y}_t | t=1, 2, \dots, n\}$ 是需求发生的自相关序列，采用自相关系数来描述这种相关性的强弱，因此可以采用两种状态的一阶马尔可夫过程进行模拟（一阶马尔可夫过程是指只考虑时间序列中 \bar{y}_t 与 \bar{y}_{t+1} 相关）。由自相关序列 $\{\bar{y}_t | t=1, 2, \dots, n\}$ 可以得到各种状态的转移概率：记 p_{00} 为连续零需求的概率， p_{01} 为零需求转为 1 的概率； p_{10} 为 1 转为零需求的概率； p_{11} 为连续非零需求的概率。

状态转移概率矩阵为

$$P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$$

利用状态转移概率矩阵，以序列中最后需求 \bar{y}_n 的值为条件，可以估计从状态 \bar{y}_n 到提前期 l ($l=1, 2, \dots, L$) 的状态转移概率。

设 $\bar{y}_n = i$ ($i=0, 1$)，则 \bar{y}_{n+1} 的预测结果为 $P(\bar{y}_{n+1}=0) = p_{i0}$ ， $P(\bar{y}_{n+1}=1) = p_{i1}$ ，即

$$[P(\bar{y}_{n+1}=0), P(\bar{y}_{n+1}=1)] = [p_{i0}, p_{i1}] \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}^{l-1} \quad (4-32)$$

设 $p_{n+1}^A = p(\bar{y}_{n+1}=1)$ ， p_{n+1}^A 为第 l 期需求为非零的概率，可以得到自相关需求序列的提前期非零需求发生概率的预测结果。

3. Logistic 回归预测

由于 $\{y_1, y_2, \dots, y_n\}$ 是二进制序列，变量的值为“0”或“1”属于分类变量而不是连续变量，不仅波动较大，而且规律性较差，不适合用线性回归判断。在分析分类变量时，通常采用对数线性模型。当对数线性模型中的某个二分类变量作为因变量并定义为一组自变量的函数时，对数线性模型就变成了 Logistic 回归模型。Logistic 回归模型是一种概率模型，它以事件的发生概率为因变量，影响因素为自变量，适用于因变量是二分类的情况。因此，在对解释变量进行回归分析时，Logistic 回归模型是一种有效的工具。

假设有 m 个解释变量 x_i ($i=1, 2, \dots, m$)，用 $\{(x'_1, x'_2, \dots, x'_m) | t=1, 2, \dots, n\}$ 表示关于 $\{y_1, y_2, \dots, y_n\}$ 的解释变量的时间序列。

根据 Logistic 函数取得为 Logistic 回归模型为

$$\hat{p}_t = \frac{1}{1 + \exp\left(-\left(a + \sum_{i=1}^m \beta_i x_i^t\right)\right)} \quad (4-33)$$

式中： \hat{p}_t 为在给定序列自变量 x_i^t 时的事件发生概率。

利用极大似然估计，可以得到参数 α 和 β 。这样，一旦得到各个时间段的解释变量构成的样本，并同时得到其事件是否发生的观测值，就能够使用这些信息来描述和分析在特定条件下事件发生的概率。通过 Logistic 回归运算，便可以得到每一个非零需求的发生概率。

算法 2 及步骤如下：

(1) 计算初始值 p_0 ($p_0 \in [0,1]$)，它使 $\sum_{t=1}^n (\text{sign}(\hat{p}_t - p_0) - y_t)$ 的值最小。其中，

$$\text{sign}(\hat{p}_t - p_0) = \begin{cases} 1 & (\hat{p}_t - p_0 \geq 0) \\ 0 & (\text{其他}) \end{cases}。$$

(2) 对每一个 \hat{p}_t ($t=1,2,\dots,n$)，如果 $\hat{p}_t \geq p_0$ ，令 $\hat{y}_t = 1$ ；否则，令 $\hat{y}_t = 0$ 。

该算法对拟合时间序列概率 $\{p'_t | t=1,2,\dots,n\}$ 设置了下限 p_0 ，将大于 p_0 的预期值变为 1，否则转化为 0，可以得到相应的备件需求发生时间序列 $\{\hat{y}_t | t=1,2,\dots,n\}$ ，这样处理的好处是使得 $\{\hat{y}_t | t=1,2,\dots,n\}$ 与实际序列的 $\{y_t | t=1,2,\dots,n\}$ 的匹配度最大。

4.4.3 需求数量预测

1. 提前期需求发生概率预测

前面已经通过 Logistic 回归预测和马尔可夫过程预测得到解释变量序列非零需求发生概率和自相关序列非零需求发生概率。对两个预测结果，通过组合算法 3，可以得到最终的提前期非零需求发生概率的预测结果。

算法 3 及步骤如下：

(1) 初始设定 r' ，令 $\tilde{y}_{n+l} = 0$ ， $l \in \{1,2,\dots,L\}$ 。

(2) 计算时间序列 $\{y_1, y_2, \dots, y_n\}$ 的自相关系数 r_1 。

(3) 如果 $|r_1| \geq r'$ ，令 $\tilde{y}_{n+l} = 0$ 的概率为 p_{n+l}^A ，转到步骤 (6)；否则，令 $l=1$ ，转到步骤 (4)。

(4) 如果 $l \leq L$, 计算 \hat{p}_{n+l} , 转到步骤 (5); 否则, 转到步骤 (6)。

(5) 如果 $\hat{p}_{n+l} \geq p_0$, 令 $\tilde{y}_{n+l} = 1$ 的概率为 \hat{p}_{n+l} ; 否则, 令 $\tilde{y}_{n+l} = 1$ 的概率为 \hat{p}_{n+l}^A ; 令 $l = l + 1$, 转到步骤 (4)。

(6) 结束。

该算法中, 步骤 (3) 表示如果需求发生时间序列 $\{y_1, y_2, \dots, y_n\}$ 具有很强的自相关性, 则采用马尔可夫过程预测提前期内非零需求的发生概率。步骤 (4) 和步骤 (5) 对预测该期需求发生概率的方法进行判别: 首先采用 Logistic 回归方法预测, 若发生非零需求的概率不小于概率下限 p_0 , 则予以确认并保留预测值; 若非零需求发生的概率低于概率下限 p_0 , 则采用两状态一阶的马尔可夫过程进行预测。然后, 对两种预测结果进行整合, 得到固定提前期备件需求发生的概率 $\{\tilde{y}_{n+1}, \tilde{y}_{n+2}, \dots, \tilde{y}_{n+l}\}$ 。

2. 改进灰自助法及仿真过程

对于只依赖于给定的观测信息, 不需要其他假设的一种新统计推断方法——Bootstrap 方法, 非常适用于备件需求的小样本估计问题。本书在灰自助动态评估模型 GBM(1,1)的基础上改进原灰自助法的抽样方法和数据拟合方法, 可以避免灰自助法在极小子样情况下的重复抽样和提高评估精度。改进灰自助法可分三步进行: Bootstrap 抽样、GM(1,1)二次数据拟合和统计预测。

1) Bootstrap 抽样

设备件需求数据样本为 $\{X = \{x(t) | t = 1, 2, \dots, n\}\}$ 。

根据 Bootstrap 抽样, 将数据从小到大排序, 得到数据序列

$$X' = \{x'(t) | t = 1, 2, \dots, n\} \quad (x'(1) \leq x'(2) \leq \dots \leq x'(n)) \quad (4-34)$$

经验分布函数值为 $0, \frac{1}{n-1}, \frac{2}{n-1}, \dots, \frac{n-2}{n-1}, 1$ 。将经验分布函数分段线性化,

则抽样过程如下: ①产生 $(0,1)$ 区间随机数 r ; ②令 $\beta = (n-1)r$, $i = \text{int}(\beta) + 1$; ③得到随机数 $X_F = X(i) + (\beta - i + 1)(X(i+1) - X(i))$ 。

按照上述方法重复抽样 m 次, 得到第一个 Bootstrap 样本。这样连续地重复 p 次得到 p 个 Bootstrap 再抽样样本, 用向量表示为

$$Y = (Y_1, Y_2, \dots, Y_b, \dots, Y_p) \quad (4-35)$$

式中: Y_b 为第 b 个 Bootstrap 样本, $Y_b = \{y_b(u) | u = 1, 2, \dots, m; b = 1, 2, \dots, p\}$, 其中, $y_b(u)$ 为 Y_b 中第 u 个 Bootstrap 再抽样样本。

2) GM(1,1)二次数据拟合

灰色系统理论要求 $y_b(u) \geq 0$, 为避免 $y_b(u) \leq 0$, 可以做如下数据平移变换:

$$X_b^{(0)} = \{y_b(u) + q\} = \{x_b^{(0)}(u)\} \quad (u=1, 2, \dots, m; \quad b=1, 2, \dots, p) \quad (4-36)$$

式中: q 为待定的参数, 只要满足 $y_b(u) \geq 0$ 即可。如果 $y_b(u) \geq 0$, 则 $q = 0$ 。
对 $X_b^{(0)}$ 做一次累加生成, 得一次累加生成序列向量为

$$X_b = \{x_b(u)\} = \left\{ \sum_{j=1}^u x_b^{(0)}(j) \right\} \quad (4-37)$$

设均值生成序列向量为

$$Z_b = \{z_b(u)\} = \int_{u-1}^u x_b(t) dt \quad (u=2, 3, \dots, m) \quad (4-38)$$

在初始条件 $x_b(1) = x_b^{(0)}(1) + c$ (c 为一个未知优化量) 下, 累加生成序列的预测值为

$$\hat{x}_0(j+1) = \left[x_b^{(0)}(1) + c - \frac{c_1}{c_2} \right] e^{-c_1 j} + \frac{c_1}{c_2} \quad (j=0, 1, \dots, m-1) \quad (4-39)$$

式中: c_1 ($c_1 \neq 0$) 和 c_2 为误差, 且 $(c_1, c_2)^T = (D^T D)^{-1} D^T (Y_b)^T$, $D = (-Z_b, I)^T$, $I = (1, 1, \dots, 1)$ 。

如果计算出 $c_1 = 0$, 则这个样本抽样失败, 需重新抽样。

为了提高模型精度, 对参数进行第二次拟合估计:

$$\hat{x}(j+1) = A e^{-c_1 j} + B(j=m-1, m) \quad (4-40)$$

式中

$$(A, B)^T = (G^T G)^{-1} G^T (X_0)^T, \quad G = (E, I), \quad I = (1, 1, \dots, 1)^T,$$

$$E = (e^0, e^{-c_1}, e^{-2c_1}, \dots, e^{-(n-1)c_1})^T, \quad X_0 = (\hat{x}_0(1), \hat{x}_0(2), \dots, \hat{x}_0(m))^T$$

根据累减生成, 就可以用仿真样本 Y_b 估计第 $w = m+1$ 次试验的数据为

$$\hat{y}_b(w) = \hat{x}_b(w) - \hat{x}_b(w-1) - q \quad (w=m+1; b=1, 2, \dots, p) \quad (4-41)$$

第 w 次试验有 p 个数据, 可构成序列向量

$$\hat{X}_w = \{\hat{y}_b(w)\} \quad (b=1, 2, \dots, p, w=m+1) \quad (4-42)$$

3) 统计预测

估计 \hat{X}_w 中的 p 个数据的分布, 用 \hat{X}_w 中数据分布去逼近 X 的分布, 进而得到备件提前期的需求分布。

3. 提前期需求分布计算

得到备件非零需求发生的概率后, 还需要预测出非零需求的具体值。
Bootstrap 方法能在过去发生的非零需求值中通过随机取样来得到未来非零需求

的预测值。因此,本书利用改进灰色自助法产生足够多的预测值,来得到提前期的需求分布,需求预测值表示为 Y_{LTD} 。

算法 4 及其步骤如下:

(1) 设定抽样次数限值为 N (抽样次数 N 通常设为 1000), 令抽样次数初始值 $C=1$ 。

(2) 若 $C \leq N$, 令 $Y_{LTD}=0$, 转到步骤 (3); 否则, 转到步骤 (5)。

(3) 对 l 从 1 到 L , 运用算法 3, 得到 \tilde{y}_{n+l} ; 若 $\tilde{y}_{n+l}=1$, 在实际需求 $d_t (t=1,2,\dots,n)$ 中以相同概率随机抽取非零需求 d'_t , 令 $\tilde{d}_{n+l}=d'_t$; 否则, 令 $\tilde{d}_{n+l}=0$ 。此时 $Y_{LTD}=Y_{LTD}+\tilde{d}_{n+l}$ 。

(4) 输出 Y_{LTD} , 令 $C=C+1$, 转到步骤 (2)。

(5) 结束。

该算法说明: 步骤 (3) 表示根据算法 3 的预测结果对提前期中的每一期产生随机数 0 或 1, 0 表示该期发生零需求, 1 表示该期发生非零需求。对于零需求, 令其需求值为 0; 对于非零需求, 采用对过去发生的非零需求值随机取样的方法对其赋值。汇总提前期内每一期的需求值, 得到一个估计值 Y_{LTD} 。通过步骤 (2)~步骤 (4) 的 N 次循环, 就可获得足够的估计值来得到提前期需求预测值的分布。

4.4.4 误差分析

对于连续性需求预测, 已经有很多方法可以用来评价预测结果的准确性, 如平均误差、平均绝对误差、均方误差、平均百分比误差、平均绝对百分比误差等, 而对于间断性需求预测结果的准确性还没有比较成型的评价方法。

现有的用于评价间断性需求预测准确度的方法主要采用连续性需求预测的准确度评价方法, 如平均百分比误差、平均绝对误差等。Syntetos 通过平均百分比误差将修正后的 Croston 方法与原方法进行比较, 说明该方法比原 Croston 方法准确。Ghobbar 采用平均绝对百分比误差来比较预测方法的预测误差。然而, 连续性需求预测结果的准确度评价方法在评价间断性需求预测结果时会遇到困难, 因为间断性需求中存在很多零值, 会导致分母为零 (如平均绝对百分比误差), 从而使得准确性难以判断。因此, 考虑从预测需求是否发生的准确度和需求量预测值大小的准确度两个方面来建立间断性需求预测结果的评价机制。

1. 非零需求发生概率预测误差分析

设待预测的第 t 期的备件实际需求量为 $X_t (t=1,2,\dots,n)$, n 为需要预测的期

数，对备件需求序列进行转换：

$$Z_t = \begin{cases} 0, & X_t = 0 \\ 1, & X_t \neq 0 \end{cases}$$

同理，设该备件第 t 期需求量的预测值为 Y_t ，并对预测需求序列进行转换：

$$Z'_t = \begin{cases} 0, & Y_t = 0 \\ 1, & Y_t \neq 0 \end{cases}$$

比较 Z'_t 与 Z_t ，得到

$$F_t = \begin{cases} 1, & Z'_t = Z_t \\ 0, & Z'_t \neq Z_t \end{cases}$$

最后通过计算非零需求预测正确率（Ratio of Occurrence Accuracy, ROA）来评价需求是否发生的预测准确度，即

$$\text{ROA} = \frac{\sum_{t=1}^n F_t}{n} \times 100\% \quad (4-43)$$

当提前期中某一期的实际需求为 0 但预测结果为 1，或实际需求为 1 而预测结果为 0 时，是一次预测错误。显然，ROA 的值越大，表明预测需求是否发生的准确度越高。

2. 需求量预测值误差分析

对于需求量预测值的误差分析，采用平均绝对百分比误差（MAPE），即

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y(t) - \hat{Y}(t)}{Y(t)} \right|$$

4.4.5 示例分析

对某军械器材仓库 2001—2010 年备件保障情况进行数据收集，选取导弹装备的 8 种重要功能部件的备件进行分析，将每半年的领用量进行累积，以此作为备件的需求量，同时将没有领用记录的半年备件需求量设置为零，从而得到以半年为时间单位的备件需求时间序列。对于这些重要功能部件，除了部件自身的可靠性之外，战备时间和通电次数对其备件的需求影响较大。这里仅以过载开关为例，2001—2010 年的装备使用与保障无量纲化处理后的数据及备件需求情况如表 4-6 所列。将 2001—2007 年的 14 期数据作为训练样本，2008—2010 年的 6 期数据作为检验样本。在组合预测方法的计算中， $\bar{r} = 0.08$ ， $r' = 0.10$ 。在 Croston 法的计算中，平滑常数 $\alpha = 0.25$ 。

表 4-6 装备使用及保障数据统计

序号	1	2	3	4	5	6	7	8	9	10
战备时间	0	0.1045	0.2386	0.5824	0.3152	0.2975	0.6791	0.4583	0.4937	0.5639
通电次数	0.7561	0.5317	0.3275	0.5879	0.2871	0.5716	0.8594	0.4387	0.3274	1
需求量	2	1	0	3	0	1	4	0	0	5
序号	11	12	13	14	15	16	17	18	19	20
战备时间	0.6484	0.7146	0.6874	0.5137	0.7847	0.4317	0.5758	1	0.4129	0.8696
通电次数	0.5972	0.9413	0.8475	0.6728	0.7249	0.4017	0.6137	0.9736	0.4635	0.8659
需求量	0	4	3	0	3	0	2	5	0	4

将导弹装备的战备时间和通电次数作为 Logistic 回归模型的解释变量。将备件需求发生转换的阈值 p_0 设为 0.50。模型参数的估计及检验采用最大似然统计分析常用的专业统计 SPSS 软件，它具有 Logistic 回归分析功能。对 Logistic 回归方程的显著性检验采用三种检验方法：总体零假设（自变量对事件发生可能性无影响作用的无效假设）的似然比检验、记分检验和 Wald 检验。由表 4-7 可知， χ^2 作用不显著，三种检验方法的 p 值均小于 0.05，因此回归模型有显著性意义。

表 4-7 检验总体零假设（ $\beta = 0$ ）

检验类型	χ^2	自由度	P
似然比检验	10.483	4	0.003
记分检验	8.671	4	0.004
Wald 检验	7.396	4	0.025

对每组样本各做抽样次数为 1000 的 Bootstrap 抽取，它们对应的自助统计量的统计特性如图 4-5 所示。从图 4-5 可以看出备件需求的频率分布近似正态分布。

图 4-6 给出了该型备件需求未来 6 期内指数平滑法、回归分析法、Croston 法、MB 法和组合预测法的预测结果。从图中可以看出：指数平滑法和回归分析法两种连续性预测方法的预测误差最大，超过了平均需求的 70%；组合预测法的预测结果比 Croston 和 MB 两种方法得到显著改善，仅在 2 期和 5 期零需求的预测上存在偏离，但预测值绝对误差仍小于 0.5。

由于 Croston 法只能预测提前期平均需求，不能预测提前期每一期的需求发生情况，无法计算 ROA，因此，只对 MB 法和组合预测法进行 ROA 比较。图 4-7 给出未来 2 期的 8 种备件需求预测结果，综合 8 种备件需求预测结果得出，组合预测法的 ROA 比 MB 方法提高了 35.6%。

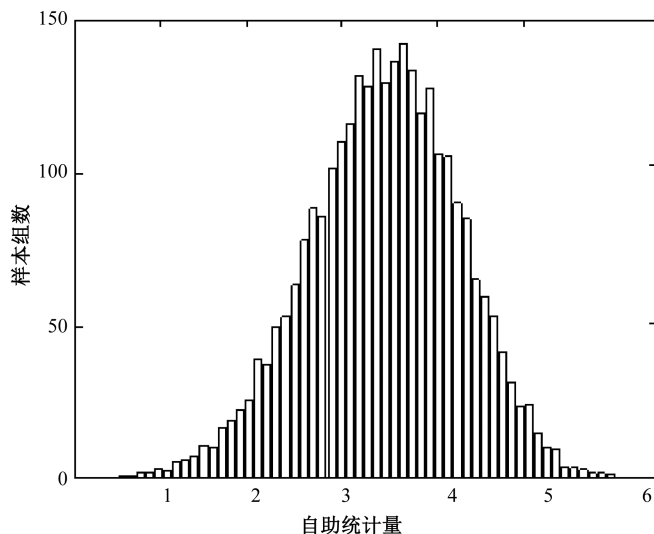


图 4-5 自助统计量直方图

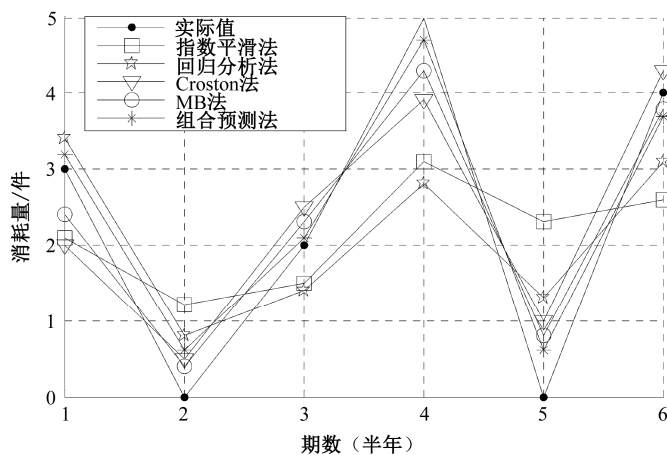


图 4-6 未来 6 期的 ROA 比较

当提前期为 2 期时，三种方法对 8 种备件预测结果的数量准确度比较如图 4-8 所示，纵坐标轴为 $1-\text{MAPE}$ ，表示需求在数量上的准确度。组合预测法对所有备件平均准确度达到 92%，远高于 Croston 法的 68.5% 和 MB 法的 79.3%。由此可见，组合预测法在 multi-period prediction 方面比 Croston 法和 MB 法具有明显的优势。

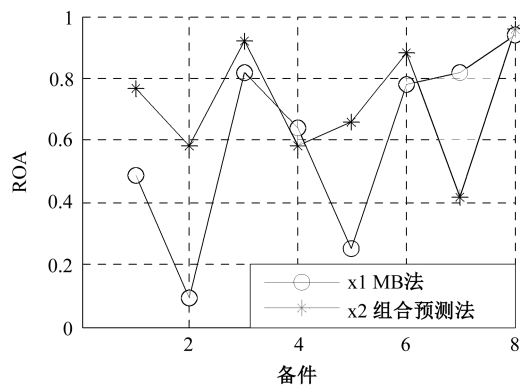


图 4-7 未来 2 期的 ROA 比较

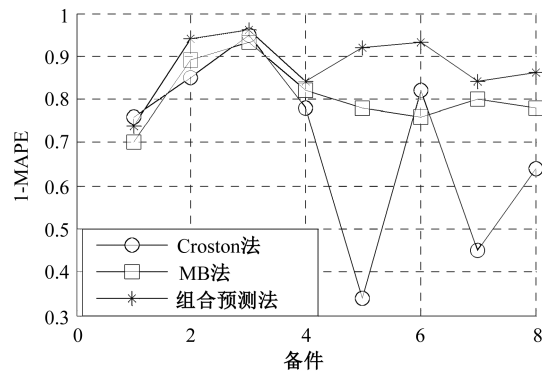


图 4-8 未来 6 期需求数量预测准确度比较

4.5 考虑设备停机的备件需求预测仿真模型

目前，武器装备备件的预测理论和方法有了很大的进展，各种新的理论和方法不断涌现，但这些方法或模型大都将备件视为在冷储备情况下，即忽略了备件在储备期间（停机期间）发生故障的可能性，而事实上导弹装备经常会出现零部件或备件在储备期间损坏而导致短缺的问题，按这样的模型预计出的结果往往与实际需求具有一定的偏差。目前，如何建立考虑设备停机的武器装备备件需求预测模型，从而更准确地确定备件需求量成为亟待解决的问题。

4.5.1 基于最大熵原理与概率加权矩的备件寿命分布确定

在建立备件需求预测模型之前,首先需要确定备件的寿命分布。武器装备故障机理十分复杂,而且其备件需求大都呈现“短周期、小样本”的特点,这就决定了采用依靠大量样本来确定备件寿命分布的常用统计分布方法,很难达到预期效果。统计经验表明,寿命分布对于常用统计分布的假定,有时在 0.05 的水平被拒绝。为了解决这一问题,本节将备件寿命分布形式看作未知,提出用最大熵法确定其寿命分布,并利用遗传算法求解建模过程中的非线性规划问题。

1. 最大熵原理

根据 Shannon 的定义,如果 x 是定义在 R 上的连续随机变量并具有概率密度函数 $f(x)$, 则 x 的熵为

$$H = -\int_R f(x) \ln f(x) dx$$

基于对信息熵本质的理解, Jaynes 指出,如果一种分布能在附加的约束条件下使得熵最大,则它是最小偏见的,这就是最大熵原理 (Maximum Entropy Principle, MEP)。具体可表示为

$$\begin{aligned} \max H = & -\max \int_R f(x) \ln f(x) dx \\ \text{s.t.} \quad & \begin{cases} \int_R f(x) dx = 1 \\ \int_R x^i f(x) dx = m_i \quad (i=1, 2, \dots, N) \\ x \geq b \text{ 或 } x < b \end{cases} \end{aligned}$$

式中: $x = G(x_1, x_2, \dots, x_n)$; i 为 x 的原点矩的阶数; b 为保证 x 有意义的量; m_i 为 x 的第 i 阶原点矩观测值,即样本的第 i 阶原点矩; $f(x)$ 为 x 的密度函数,模型的约束条件为 $m+2$ 个。

基于最大熵原理,推断定义 x 的最小偏见的分布模型如下:

$$\begin{aligned} \max H = & -\int_R f(x) \ln f(x) dx \\ \text{s.t.} \quad & \int_R x^n f(x) dx = v_n \quad (n=0, 1, 2, \dots, N) \end{aligned} \quad (4-44)$$

式中: n 为数据样本量; N 为样本矩的最高阶数; x_i 为第 i 个随机变量; v_n 为原点矩。

经典的样本矩定义为

$$\nu_n = \frac{1}{m} \sum_{i=1}^m x_i^n$$

或

$$c_n = \frac{1}{m} \sum_{i=1}^m (x_i - \nu_1)^n$$

式中： c_n 为 n 阶中心矩。原点矩 ν_n 与中心矩 c_n 可以通过二项式定理实现相互转换。

对于上述模型，引入拉格朗日算子，采用经典的变分法解得最大熵概率密度函数解析式为

$$f(x, \lambda_0, \lambda_1, \dots, \lambda_n) = \exp \left(\lambda_0 + \sum_{n=1}^N \lambda_n x^n \right) \quad (4-45)$$

式中： $\lambda_i (i=0, 1, 2, \dots, N)$ 为与第 i 阶矩约束对应的拉格朗日算子。

根据式 (4-45)，只要确定参数 $\lambda_i (i=0, 1, 2, \dots, N)$ 就可以完全确定概率密度函数 $f(x)$ ，然而，基于不同样本量估计的模型对于系统统计特性的描述偏差相差极大。为克服此缺陷，引入概率加权矩方法。

2. 概率加权矩估计方法

1) 概率加权矩

定义在 R 上的连续随机变量 x ，若其累积分布函数 $F(x) = P(X \leq x)$ 及其逆函数 $x = x(F)$ 存在，则其概率加权矩 (Probability Weighted Moments, PWM) 为

$$M_{i,n,k} = \int_0^1 [x(F)]^i F^n (1-F)^k dF \quad (4-46)$$

式中：下标变量 i, n, k 为实数。

由式 (4-46) 不难看出，概率加权矩是经典矩的推广，而经典矩则是概率加权矩的特例。

简单而常用的两类概率加权矩为

$$\alpha_k = M_{1,0,k} = \int_0^1 x(F) (1-F)^k dF \quad (4-47)$$

$$\beta_n = M_{1,n,0} = \int_0^1 x(F) F^n dF \quad (4-48)$$

对抽样得来的数据进行排序，得到 $x_1 \leq x_2 \leq \dots \leq x_{m-1} \leq x_m$ ，则 α_k 与 β_n 的无偏估计 a_k 和 b_n 分别为

$$a_k = \frac{1}{m} \sum_{i=1}^m \left[\binom{m-1}{k} x_i \right] / \binom{m-1}{k} \quad (k=0,1,2,\dots,m-1) \quad (4-49)$$

$$b_n = \frac{1}{m} \sum_{i=1}^m \left[\binom{i-1}{n} x_i \right] / \binom{m-1}{n} \quad (n=0,1,2,\dots,m-1) \quad (4-50)$$

式中

$$\binom{i}{n} = \frac{i!}{(i-n)!n!}$$

由于基于概率加权矩的参数估计方法对小样本量没有经典矩方法敏感,故概率加权矩方法具有较好的鲁棒(Robust)特性。与经典矩的统计相比,概率加权矩克服了在小样本量下经典矩估计常常出现的大偏差问题。

2) 概率加权矩与逆累积分布函数

对于非负随机变量 x , β_n 可以看作其逆累积分布函数的矩。这一结论可由下式得到:

$$E(X^n) = \int_R x^n f(x) dx = \int_0^1 [x(u)]^n du \quad (4-51)$$

式中

$$du = dF(x) = \frac{f(x)}{\int f(x) dx}$$

比较 β_n 关系式与式(4-51)可知, $x(u)$ 和 u 分别相似于 $f(x)$ 和 x 。Mallows 和 Kadane 引入了下述变换关系式:

$$dT(u) = \frac{x(u)du}{\int_0^1 x(u)du} = \frac{x(u)du}{\beta_0} \quad (4-52)$$

基于 dT 可得概率加权矩 β_n 另外一种形式为

$$\beta_n = \beta_0 \int_0^1 u^n dT(u) \quad (n=1,2,\dots,N) \quad (4-53)$$

式中: β_0 为样本均值。 β_n/β_0 为逆累积分布函数 $x(u)$ 的 n 阶矩,与此类似, α_n/α_0 为逆累积分布函数 $x(1-u)$ 的 n 阶矩。对于 $\beta_0=1$ 的标准化形式, β_n 为逆累积分布函数 $x(u)$ 的 n 阶矩。

综上所述,可得如下最大熵逆累积分布函数模型:

$$\begin{aligned} \max H &= -\int_0^1 x(u) \ln[x(u)] du \\ \text{s.t. } \int_0^1 u^n x(u) du &= b_n \quad (n=0,1,2,\dots,N) \end{aligned} \quad (4-54)$$

解式(4-54)可得

$$x(u, \lambda_0, \lambda_1, \dots, \lambda_n) = \exp\left(\lambda_0 + \sum_{n=1}^N \lambda_n u^n\right) \quad (4-55)$$

对比式(4-44)与式(4-54)可以发现,若定义域做出相应改变,式(4-54)中约束等于各阶经典矩,则式(4-55)为概率密度函数,约束等于各阶概率加权矩,则式(4-55)为逆累积分布函数。

3. 参数估计

下面首先确定乘子 $\lambda_i (i=0,1,2,\dots,N)$ 的计算公式,然后给出算法实现步骤。

1) 计算公式

由逆累积分布函数的零阶矩等于样本的平均值,可得

$$\int_0^1 \exp\left(\lambda_0 + \sum_{n=1}^N \lambda_n u^n\right) du = b_0 \quad (4-56)$$

两边同乘 $e^{-\lambda_0}$, 可得

$$\int_0^1 \exp\left(\sum_{n=1}^N \lambda_n u^n\right) du = e^{-\lambda_0} \cdot b_0 \quad (4-57)$$

或

$$\lambda_0 = -\ln\left[\int_0^1 \exp\left(\sum_{n=1}^N \lambda_n u^n\right) du\right] + \ln b_0 \quad (4-58)$$

对式(4-57)中 λ_n 进行微分,可得

$$-e^{-\lambda_0} \cdot b_0 \cdot \frac{\partial \lambda_0}{\partial \lambda_n} = \int_0^1 u^n \exp\left(\sum_{n=1}^N \lambda_n u^n\right) du \quad (4-59)$$

或

$$\frac{\partial \lambda_0}{\partial \lambda_n} = -\frac{1}{b_0} \int_0^1 u^n \exp\left(\lambda_0 + \sum_{n=1}^N \lambda_n u^n\right) du \quad (4-60)$$

将式(4-55)代入式(4-60),并结合式(4-54)可得

$$\frac{\partial \lambda_0}{\partial \lambda_n} = -\frac{b_n}{b_0} \quad (4-61)$$

对式 (4-61) 中 λ_n 进行微分, 可得

$$\frac{\partial \lambda_0}{\partial \lambda_n} = -\frac{\int_0^1 u^n \exp\left(\sum_{n=1}^N \lambda_n u^n\right) du}{\int_0^1 \exp\left(\sum_{n=1}^N \lambda_n u^n\right) du} \quad (4-62)$$

由式 (4-61) 和式 (4-62) 可得

$$b_n = b_0 \cdot \frac{\int_0^1 u^n \exp\left(\sum_{n=1}^N \lambda_n u^n\right) du}{\int_0^1 \exp\left(\sum_{n=1}^N \lambda_n u^n\right) du} \quad (4-63)$$

以上是关于 $\lambda_i (i=0,1,2,\dots,N)$ 的 N 个方程。

为便于数值求解, 将式 (4-63) 改写为

$$Q_n = 1 - b_0 \cdot \frac{\int_0^1 u^n \exp\left(\sum_{n=1}^N \lambda_n u^n\right) du}{b_n \cdot \int_0^1 \exp\left(\sum_{n=1}^N \lambda_n u^n\right) du} \quad (4-64)$$

确定如下优化问题:

$$\min Q = \sum_{n=1}^N Q_n^2 \quad (4-65)$$

应用遗传算法 (Genetic Algorithm, GA) 求解式 (4-65), 可得各个 $\lambda_i (i=0,1,2,\dots,N)$ 。

上述结果是可靠的, 下述两个定理保证了上述方法求解结果的存在性与唯一性。

定理 4.1 式 (4-44) 最大熵解存在且唯一的充分必要条件是给定的矩 $\beta_n (n=0,1,2,\dots,m-1)$ 是完全单调序列 $\beta_n (n=0,1,2,\dots)$ 的一部分。

定理 4.2 最大熵解 $x(u)$ 平均值收敛于 $\bar{x}(u)$, 即对于 $[a,b]$ 上的连续函数 $X(u)$, $X(u)$ 对于 $x(u)$ 的矩收敛于 $X(u)$ 对于 $\bar{x}(u)$ 的矩, 即

$$\lim_{N \rightarrow \infty} \int_R X(u) x(u) du = \lim_{N \rightarrow \infty} \int_0^1 X(u) x(u) du = \int_0^1 X(u) \bar{x}(u) du \quad (4-66)$$

2) 算法步骤

(1) 对样本时间序列进行排序;

(2) 根据已知排序样本数据计算各阶概率加权矩, 为满足一致性条件一般最多只需用到 6 阶矩;

(3) 建立残差优化模型式 (4-65), 确定优化初始值;

(4) 调用遗传优化算法子程序;

(5) 判断是否收敛, 若收敛则转步骤 (7), 否则转步骤 (6);

(6) 用另一初始值重新计算, 转步骤 (4);

(7) 输出 $\lambda_i (i = 0, 1, 2, \dots, N)$ 的值;

(8) 计算结束。

4.5.2 基于保障度的间断工作备件需求预测模型建立

1. 备件供应保障过程描述

假设某型武器装备备件供应保障系统由 $S+1$ 个单元组成, 如图 4-9 所示。如果武器装备处于工作状态, 其中 1 个单元工作, 其余没有发生故障的单元都做温储备, 当工作单元故障后, 由其下一个没有发生故障的储备单元替换; 如果武器装备处于停机状态, 所有未发生故障的单元均做温储备。

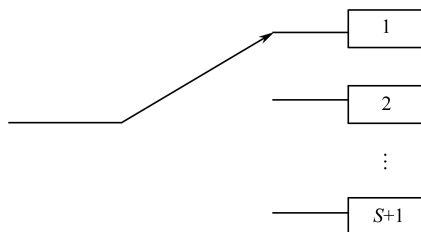


图 4-9 温储备系统

2. 建模条件设定

为建立武器装备备件需求预测模型, 做如下假设:

(1) 武器装备总量及单装零部件数量一定。

(2) 开始时刻, 武器装备系统所有零部件及备件均为新品, 其中武器装备上 1 个零部件在工作, S 个备件储备起来用于替换武器装备上损坏的零部件。

(3) 在每个阶段状态中, 系统部件和备件都只有正常与故障两种状态, 且相同零部件或备件寿命分布相同。

(4) 武器装备及其部件只有“工作”或“不工作”两种状态,即备件供应保障周期为 1 年(365 天),导弹装备每周工作 t_w (h),不工作时间 $t_s = 168 - t_w$ (h)。武器装备工作期间,安装在武器装备上的零部件处于工作状态,其余备件处于储备状态。武器装备停机期间,零部件的故障率与备件的故障率相同。

(5) 零部件及备件均不可修,武器装备故障时通过更换备件予以排除,且用备件更换故障部件后功能不下降,换件时间忽略不计。

(6) 如果武器装备在工作期间发生故障,则立即维修;如果武器装备在停机期间故障,则等到再次开机(或检测)时进行维修。

(7) 武器装备系统各组成部件的失效相互独立,其失效不会发生在同一时刻。

(8) 任一部件在发现其不能工作之前总是完好的,即不能工作时间从故障发现时开始。

3. 模型建立

基于上述假设,零部件在时间区间 $[0, t](t \leq t_w)$ 内累计故障概率为 $\int_0^t \lambda e^{-\lambda t} dt$,

利用累计故障概率相等原则,即累计损伤相等原则,把故障率为 λ 的零部件在时间区间 $[0, t]$ 内的累计故障概率折合到故障率为 μ 的概率分布上,则

$$\int_0^t \lambda e^{-\lambda t} dt = \int_0^{t'} \mu e^{-\mu t} dt$$

可推导出

$$t' = \frac{\lambda}{\mu} t \quad (4-67)$$

即某零部件在故障率为 λ 的条件下工作到 t 时刻,累计故障概率与在故障率为 μ 的条件下储备到 t' 时刻相同。同理,备件在时间区间 $[0, t']$ 内的累计故障概率与零部件在时间区间 $\left[0, \frac{\mu}{\lambda} t'\right]$ 内的累计故障概率相同。

根据前面假设,在不发生故障的前提下,第 1 个零部件从初始时刻工作到 t_w 时刻,利用累计故障概率相等原则,折合到故障率为 μ 的概率分布上,相当于储备时间为 $\frac{\lambda}{\mu} t_w$;接着导弹装备停机 t_s ,储备时间变成 $\frac{\lambda}{\mu} t_w + t_s$,折合成工作时间为 $\frac{\mu}{\lambda} \left(\frac{\lambda}{\mu} t_w + t_s \right)$;然后导弹装备又工作 t_w ,工作时间变成 $\frac{\mu}{\lambda} \left(\frac{\lambda}{\mu} t_w + t_s \right) + t_w$,折

合成储备时间为 $\frac{\lambda}{\mu} \left[\frac{\mu}{\lambda} \left(\frac{\lambda}{\mu} t_w + t_s \right) + t_w \right]$ ；紧接着设备又停机 t_s ，储备时间变成 $\frac{\lambda}{\mu} \left[\frac{\mu}{\lambda} \left(\frac{\lambda}{\mu} t_w + t_s \right) + t_w \right] + t_s$ ；依此类推，直至零部件发生故障，可以推导出第 1 个零部件的累计故障分布函数，即

$$F_1(t_1) = \int_0^{t_w} \lambda e^{-\lambda t} dt + \int_{\frac{\lambda}{\mu} t_w}^{\frac{\mu}{\lambda} t_w + t_s} \mu e^{-\mu t} dt + \int_{\frac{\mu}{\lambda} \left(\frac{\lambda}{\mu} t_w + t_s \right)}^{\frac{\mu}{\lambda} \left(\frac{\lambda}{\mu} t_w + t_s \right) + t_w} \lambda e^{-\lambda t} dt + \int_{\frac{\lambda}{\mu} \left[\frac{\mu}{\lambda} \left(\frac{\lambda}{\mu} t_w + t_s \right) + t_w \right]}^{\frac{\lambda}{\mu} \left[\frac{\mu}{\lambda} \left(\frac{\lambda}{\mu} t_w + t_s \right) + t_w \right] + t_s} \mu e^{-\mu t} dt + \dots$$

通过推导可得

$$F_1(t_1) = \sum_{i=0}^{n_1} \left(\int_{it_w + it_s}^{(i+1)t_w + it_s} \lambda e^{-\lambda t} dt + \int_{(i+1)t_w}^{(i+1)t_w + (i+1)t_s} \mu e^{-\mu t} dt \right) + \int_{(n_1+1)t_w}^{(n_1+1)t_w + (n_1+1)t_s} \lambda e^{-\lambda t} dt \quad (4-68)$$

式中： i 为天数； t_1 为第 1 个零部件故障前最后一次开机保持良好状态的时间。

第 1 个零部件的寿命为

$$T_1 = 24n_1 + t_1 \quad (4-69)$$

第 1 个零部件服役期间，第 2 个备件一直处于储备状态，如果储备期间没有发生故障，第 1 个零部件故障后则用其替换，第 2 个备件的累计故障分布函数为

$$F_2(t_2) = \int_0^{T_1} \mu e^{-\mu t} dt + \int_{T_1}^{T_1 + \frac{\mu}{\lambda} t_w - t_1} \lambda e^{-\lambda t} dt + \int_{T_1 + (t_w - t_1) \frac{\lambda}{\mu}}^{T_1 + (t_w - t_1) \frac{\lambda}{\mu} + t_s} \mu e^{-\mu t} dt + \int_{(T_1 + t_s) \frac{\mu}{\lambda} + 2t_w - t_1}^{(T_1 + t_s) \frac{\mu}{\lambda} + 2t_w - t_1} \lambda e^{-\lambda t} dt + \dots$$

$$\int_{(T_1 + 2t_s) \frac{\mu}{\lambda} + (2t_w - t_1) \frac{\lambda}{\mu}}^{(T_1 + 2t_s) \frac{\mu}{\lambda} + (2t_w - t_1) \frac{\lambda}{\mu}} \lambda e^{-\lambda t} dt + \dots$$

通过推导可得

$$F_2(t_2) = \int_0^{T_1} \mu e^{-\mu t} dt + \int_{T_1}^{T_1 + \frac{\mu}{\lambda} t_w - t_1} \lambda e^{-\lambda t} dt + \sum_{i=1}^{n_2} \left(\int_{T_1 + (it_w - t_1) \frac{\lambda}{\mu}}^{T_1 + (it_w - t_1) \frac{\lambda}{\mu} + it_s} \mu e^{-\mu t} dt + \int_{(T_1 + it_s) \frac{\mu}{\lambda} + (i+1)t_w - t_1}^{(T_1 + it_s) \frac{\mu}{\lambda} + (i+1)t_w - t_1} \lambda e^{-\lambda t} dt \right) + \dots$$

$$\int_{(T_1 + (n_2+1)t_s) \frac{\mu}{\lambda} + (n_2+1)t_w - t_1 + t_2}^{(T_1 + (n_2+1)t_s) \frac{\mu}{\lambda} + (n_2+1)t_w - t_1 + t_2} \lambda e^{-\lambda t} dt$$

(4-70)

寿命 $T_2 = 24(n_1 + n_2) + t_2$ 。当第 2 个零部件故障后,接着用第 3 个备件替换,依此类推,第 N 个备件的累计故障分布函数为

$$\begin{aligned}
 F_N(t_N) = & \int_0^{T_{N-1}} \mu e^{-\mu t} dt + \int_{T_{N-1} \frac{\mu}{\lambda}}^{T_{N-1} \frac{\mu}{\lambda} + t_w - t_{N-1}} \lambda e^{-\lambda t} dt + \int_{T_{N-1} + (t_w - t_{N-1}) \frac{\lambda}{\mu}}^{T_{N-1} + (t_w - t_{N-1}) \frac{\lambda}{\mu} + t_s} \mu e^{-\mu t} dt + \\
 & \int_{(T_{N-1} + t_s) \frac{\mu}{\lambda} + 2t_w - t_{N-1}}^{(T_{N-1} + 2t_s) + (2t_w - t_{N-1}) \frac{\lambda}{\mu}} \lambda e^{-\lambda t} dt + \int_{(T_{N-1} + t_s) + (2t_w - t_{N-1}) \frac{\lambda}{\mu}}^{(T_{N-1} + t_s) + (2t_w - t_{N-1}) \frac{\lambda}{\mu} + t_s} \mu e^{-\mu t} dt + \dots
 \end{aligned} \quad (4-71)$$

迭代累计计算单元的寿命 $T_N = T_{N-1} - (t_w - t_{N-1}) + 24n + t_N$, 其中, t_N 表示单元最后一次开机后良好运行的时间。

这样依次进行备件更换,直至没有备件替换,就认为导弹装备处于不能工作状态,则最后一个没有在储备期间发生故障的零部件寿命是备件的可保障时间。

4.5.3 仿真思路及流程规划

1. 蒙特卡罗统计实验模型

通常认为任一设备的寿命都是一个服从某一分布的随机变量。根据蒙特卡罗直接抽样法,在假定设备可靠性服从指数分布时,可利用设备的平均故障间隔时间 (MTBF),通过抽取 $[0,1]$ 内的均匀随机数 ξ ,反求可靠度 (R) 函数来预测设备的随机寿命 $T(T = -MTBF \times \ln \xi)$ 。

2. 仿真思路

先得到每个单元在其寿命期内的累计故障概率抽样值,用 η_j 表示, η_j 为 $[0,1]$ 区间均匀分布的随机数 ($j=1,2,\dots,S+1$)。在第 1 个单元未发生故障以前,其余 S 个单元均处于温储备状态,也可能发生故障。因此,当第 1 个单元在工作期间发生故障需进行换件维修时,首先要判断第 2 个单元是否已经发生故障。如果没有故障,则立即进行换件维修;如果第 2 个单元在第 1 个单元发生故障前就已经故障,那么按以上方法对第 3 个单元进行判断。依此类推,直到没有备件可以更换,就认为设备故障。当最后一个替换单元的寿命大于等于备件保障周期时,就认为备件的数量满足需要;否则,认为不满足需要。按照蒙特卡罗思想,将以上过程重复 N_{mn} 次,若统计其中备件满足需要的次数为 N_{suc} ,当 N_{mn} 足够大时,

$N_{\text{suc}}/N_{\text{mn}}$ 为备件的保障概率。针对可靠度、故障概率和系统平均寿命的计算，进行 1000 次仿真运行即可，取 $N_{\text{mn}}=1000$ 次。仿真流程如图 4-10 所示。

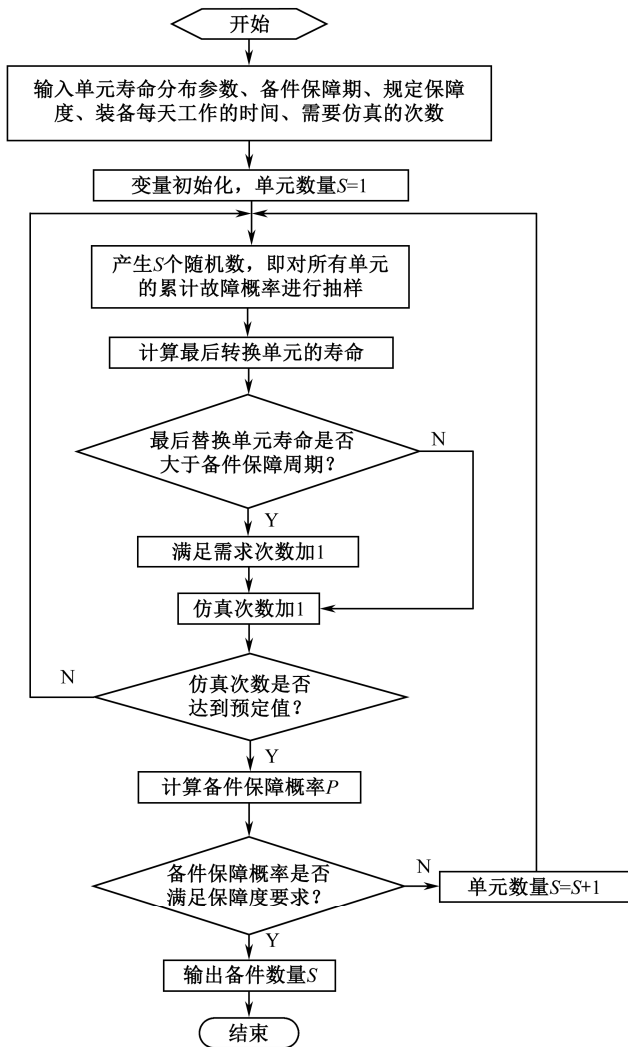


图 4-10 仿真流程

3. 仿真步骤

(1) 输入单元的工作寿命分布参数 λ 、储备寿命分布参数 μ 、模拟次数 N_{mn} 、每周工作时间 t_w 、规定保障度 P_0 。

(2) 对变量进行初始化: 已仿真次数 $N_{\text{em}} = 1$, 单元寿命 $T = 0$, 备件需求量 $N = 0$, 保障概率 $P = 0$, 单元数量 $S = 1$ 。

(3) 由蒙特卡罗直接抽样法, 产生 S 个 $[0, 1]$ 之间均匀分布的随机数, 分别表示第 $1 \sim S$ 个单元整个寿命期内的累计故障分布概率, 即 $F_j(t_j) = \eta_j (j = 1, 2, \dots, S+1)$, 确定最先发生的故障事件。

(4) 判断 N 与 S 的大小。当 $N < S$ 时, 转到步骤 (5); 当 $N = S$ 时, 转到步骤 (7)。

(5) 计算单元寿命, 如果

$$\eta_N = F_N(t_N) \leq \int_0^{T_{N-1}} \mu e^{-\mu t} dt$$

就说明该单元在储备期已经故障, 则 $N = N + 1$, 继续下一个单元的判断。如果

$$\eta_N = F_N(t_N) > \int_0^{T_{N-1}} \mu e^{-\mu t} dt$$

则说明该单元在储备期还没有故障, 可以用来替换设备上已损坏的单元, 通过式 (4-69) 计算单元的寿命时间。如果 $t_{N-1} \geq t_w$, 则说明单元是在停机期间发生故障; 如果 $t_{N-1} < t_w$, 则说明单元在工作过程中发生故障。

(6) $N = N + 1$, 转到步骤 (4)。

(7) 判断 T_N 与 $364 \times 24 + t_w$ 的大小。如果 $T_N \geq 364 \times 24 + t_w$, 说明备件数量满足需要, 则满足需要的次数 $N_{\text{suc}} = N_{\text{suc}} + 1$; 否则, 说明备件数量不满足需要。

(8) 判断仿真进行次数 N_{em} 与需要仿真次数 N_{mn} 的大小。当 $N_{\text{em}} < N_{\text{mn}}$ 时, $N_{\text{em}} = N_{\text{em}} + 1$, 转到步骤 (3); 当 $N_{\text{em}} = N_{\text{mn}}$ 时, 计算备件保障概率 $P = N_{\text{suc}} / N_{\text{mn}}$ 。

(9) 判断 P 与规定保障度 P_0 的大小。当 $P < P_0$ 时, 则 $S = S + 1$, 转到步骤 (3); 当 $P \geq P_0$ 时, 转到步骤 (10)。

(10) 输出备件数量 S , 仿真结束。

4.5.4 示例分析

某型导弹检测设备是一套以工业计算机控制的综合检测设备, 用于完成四型机载导弹的检测。该型检测设备平时主要处于停机与使用两种状态, 即存放与工作两种模式, 且检测设备和备件的储存条件基本相同。检测设备发生故障后采用换件维修的方式。该设备的某型激励通道插件板是使用最为频繁的部件之一, 统计分析其工作和储备期间的故障情况, 如表 4-8 所列。结合部队战备任务情况, 要求备件保障度为 0.9。

表 4-8 故障统计

名称	工作期间故障间隔时间/h					储存期间故障间隔时间/h				
某型插件板	350	103	253	116	189	1721	17150	5168	4374	11793
	128	231	364	282	486	7614	15466	3696	1478	2548
	125	474	227	592	171	22030	15281	4477	3840	15709
	148	233	346	283	192	3567	4037	12776	1565	8412

1. 参数确定

该插件板故障数据样本量小, 适合采用基于最大熵原理和概率加权矩的方法确定其寿命分布。所用矩的阶数 m 取 2。

在构造 GA 时, 由于目标函数是求最小值, 同时为增大适应度函数间的差距, 以提高优秀个体被选中的概率, 因此选择目标函数值倒数的 10000000000 倍作为 GA 的适应度函数, 求得其参数为 $\lambda_0 = 6.25 \times 10^{-3} \text{ h}^{-1}$, $\lambda_1 = 4.49 \times 10^{-3} \text{ h}^{-1}$, $\lambda_2 = 2.63 \times 10^{-3} \text{ h}^{-1}$ 。

从图 4-11 可以看出, 优化效果比较理想。经拟合优度检验知, 该插件板工作寿命服从参数 $\lambda = 6.25 \times 10^{-3} \text{ h}^{-1}$ 的指数分布。作为对比, 在不同样本量下, 分别计算基于概率加权矩与基于经典矩对于理论总体均值的相对误差, 计算结果如图 4-12 所示。计算结果表明, 样本量低于 10 时, 采用经典矩方法相对误差达到 50.0337%, 而采用概率加权矩方法估计的相对误差仅为 3.9694%。此结果显示: 若样本量小于或等于 30 时, 不宜采用经典矩方法。同理, 可求得储存寿命服从参数 $\mu = 1.8518 \times 10^{-4} \text{ h}^{-1}$ 的指数分布。

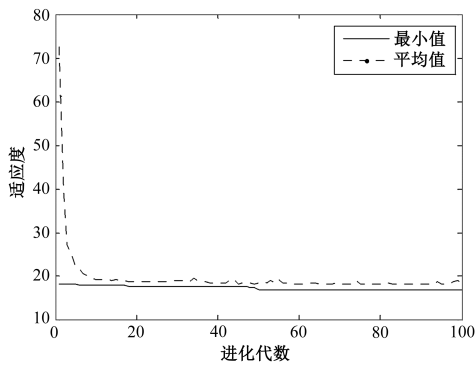


图 4-11 适应度函数图

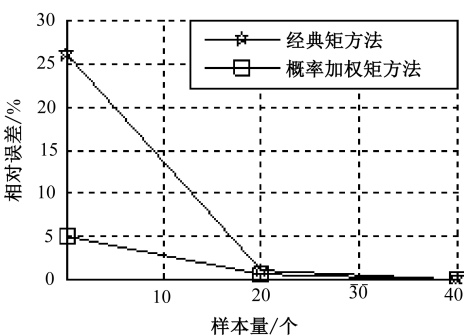


图 4-12 不同样本下两种方法计算相对误差

2. 需求仿真

利用 Matlab 编程进行 1000 次仿真, 该插件板 1 年 (52 周) 的备件需求量部分计算结果如表 4-9 所列。其中, 该检测设备每周工作 20h 时, 1 年所需备件需求量如图 4-13 所示, 备件需求量与备件保障概率的关系如图 4-14 所示。由此可以看出, 当备件数量为 4 时, 备件的满足率就达到了 92% 以上, 与表 4-9 结果相符。

表 4-9 备件需求量部分计算结果

t_w / h	0	4	8	12	16	20	48	96	168
备件需求量 $S/\text{个}$	1	2	3	3	3	4	4	5	8
实际保障概率 P	0.9324	0.9105	0.9328	0.9149	0.9212	0.9251	0.9026	0.8733	0.7469

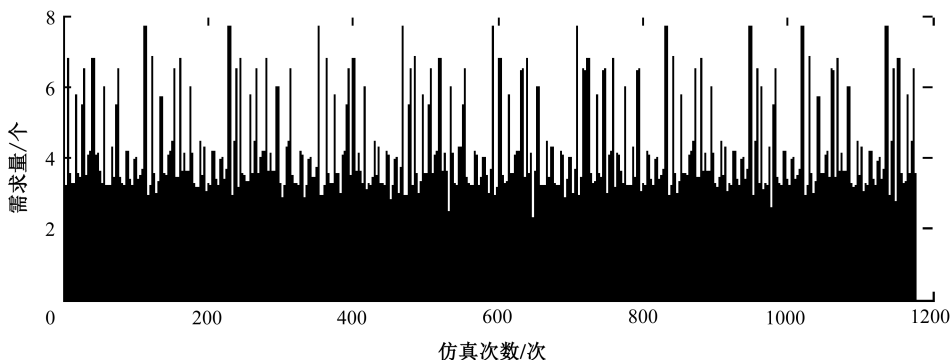


图 4-13 备件需求量仿真图

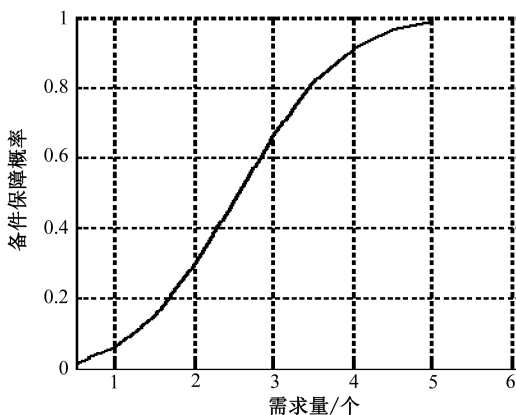


图 4-14 备件需求量与保障概率的关系曲线

3. 结果分析与评价

表 4-9 中, $t_w = 0$ 表示备件纯储备 1 年, 可视为温储备问题, 即所有单元均以 $\mu = 1.8518 \times 10^{-4} \text{ h}^{-1}$ 的故障率进行温储备, 可利用并联系统公式进行计算, 结果为 $S = 1$, $P = 0.9286$ 。 $t_w = 168$ 可视为连续工作的热储备系统, 利用热储备系统公式可解得 $S = 8$, $P = 0.7637$ 。可见, 利用仿真方法与利用解析法所得结果基本吻合, 说明本节仿真方法是可行的。

基于仿真的备件需求预测方法能够充分考虑各种影响因素, 在最贴近实际的情况下预测备件的需求量, 适用于备件需求存在随机波动的情况。

第 5 章

基于时间序列和回归分析的 武器装备备件需求预测方法



在武器装备使用阶段中后期，装备从磨合期进入偶然故障期，装备故障率较低且基本恒定。由于装备故障主要是由偶然因素引起的，因此，装备零部件因自身故障而导致的消耗会相应减少。装备形成战斗力后，部队的训练也常规化，演习和战备任务会加重，再加上一些环境和人为等因素，这都会增加备件消耗量，这一时期的备件消耗量往往具有一定的波动性，但总的趋势相对稳定。在这个阶段，通过装备使用与保障工作的开展，积累了一些与备件有关的保障数据，但是还不够充足。对于这一时期的备件需求预测，以备件历史消耗数据作为依据的时间序列预测方法、回归分析法比较常用。

5.1 时间序列法概述

时间序列分析方法是通过对过去观察值组成的序列样本的分析研究，找出动态过程的特性、最佳数学模型、估计模型参数，并利用数学模型对统计预测进行精度检验。它以时间为独立变量，利用时序模型不需要知道影响预测变量的因果关系，在系统动态性较强、影响预测变量的决定性因素少时可以起到事半功倍之效。时间序列分析主要有两个目的：一是提取时间序列的优势周期；二是构造拟

合序列的最佳数学模型，并用此模型对未来进行预报。

时间序列分析方法的基本思想是根据系统的有限长度的运行记录，建立能够比较精确地反映时间序列中所包含的动态依存关系的数学模型，并借以对系统的未来行为进行预报。时间序列的特点往往表现在四个方面：①长期趋势的变化 X_t ，序列随时间呈现的倾向性变化；②季节周期性变化 S_t ，序列随季节呈现有规律的周期变化；③循环变化 C_t ，序列以不固定的周期呈现出的波动性变化；④随机变化 ε_t ，序列在各种不确定因素作用下的无规则变化。

时间序列模型可分为加法模型和比例模型：

加法模型为

$$\gamma_t = X_t + S_t + C_t + \varepsilon_t \quad (5-1)$$

比例模型为

$$\gamma_t = X_t \cdot S_t \cdot C_t \cdot \varepsilon_t \quad (5-2)$$

时间序列法应用普遍，时间序列法要求时间序列具有稳定的变化趋势，但对于不确定数据或者间断数据等外界的干扰因素较多时，预测的可信度较差。

5.2 基于时间序列的备件需求预测方法

5.2.1 算数平均预测法

算数平均预测法是一种非常简单的方便的方法，它是根据历史上备件消耗中发生的事实，求平均值作为今后的预测值。

1. 简单算数平均法

这种方法可计算某研究对象在过去 n 个周期的实际动态数据的算数平均值，以其作为以后周期的预测值。其计算公式为

$$\hat{y} = \bar{y} = \sum_{i=1}^n y_i / n \quad (5-3)$$

式中： \hat{y} 为预测值； $y_i (i=1, 2, 3, \dots, n)$ 为历史周期实际消耗量。

2. 加权算数平均法

在简单的算数平均法中，认为过去每个周期的需求量对后续周期的需求量的

影响程度是一致的,但事实往往并非如此。因此,根据各个周期的重要程度赋以权值就可以相应提高预测精度。其计算公式为

$$\hat{y} = \sum_{i=1}^n w_i y_i \quad (5-4)$$

式中: w_i 为各历史周期的权值,且 $0 \leq w_i \leq 1$, $\sum_{i=1}^n w_i = 1$ 。

通常情况下,距离预测周期越近的数据对预测周期影响越大,因此,一般近期权值应取较大的值。算数平均预测法以平均值作为今后的预测值,适用于平时需求量变化不大比较平稳的某型备件,做近期和短期预测。

5.2.2 移动平均预测法

在武器装备备件的实际保障中,各个周期的需求量往往具有很大的随机性,为了弱化这种随机性,减小预测风险,可以采用移动平均预测法。

1. 一次移动平均法

一次移动平均法是以一组观察序列的平均值作为下一周期的预测值的方法。但是,它要不断地根据得到的统计数据“吐故纳新”,增加新近数据除去远期数据,逐渐向前移动,计算出新的平均值。其计算公式为

$$M_t^{(1)} = \frac{y_t + y_{t-1} + \cdots + y_{t-N+1}}{N} \quad (5-5)$$

式中: t 为时间序列号; $M_t^{(1)}$ 为 t 时刻的一次移动平均预测值; y_t 为 t 时刻的实际值; N 为计算一次移动平均数所选的数据个数。

2. 加权移动平均法

一次移动平均法是对参加滑动平均的数据一律平均对待,不分先后,而实际上,最新的数据往往更能反映消耗趋势。因此,要特别强调新数据的影响,突出新数据的作用。为达到此目的,可采用加权移动平均法,给予新数据较大的权值。其计算公式为

$$\bar{M}_t^{(1)} = (\alpha_0 y_t + \alpha_1 y_{t-1} + \cdots + \alpha_{N-1} y_{t-N+1}) / N = \frac{1}{N} \sum_{j=0}^{N-1} \alpha_j y_{t-j} \quad (5-6)$$

式中: α_j 为加权因子,且满足 $\sum_{j=1}^{N-1} \alpha_j = N$ 。

采用加权移动平均法可以更准确地反映实际情况。这种方法所加的“权”和“平均时段”的长短都是重要的参数，应当注意的是：移动时段长短的选择带有一定的经验性，随着 N 值的增大，预测结果越平稳，对实际消耗量的真实变化的反映就越迟钝，因此 N 值大或小各有利弊；而权值的选择也具有一定的主观性，新数据的权值越大，其预测风险就越大，越容易受随机干扰的影响。上述两个参数值的选择将很大程度上影响预测结构的精确度（对序列的了解掌握程度非常重要）。

3. 二次移动平均法

在备件需求序列中会遇到线性增长趋势，可以采用二次移动平均法。二次移动平均法是在一次移动平均法的基础上再次进行移动求取平均值的算法，但往往并不是直接计算二次移动的平均值，而是建立线性模型。其计算公式为

$$M_t^{(2)} = \frac{M_t^{(1)} + M_{t-1}^{(1)} + \cdots + M_{t-N+1}^{(1)}}{N} \quad (5-7)$$

式中： $M_t^{(1)}$ 为 t 时刻的一次平均移动数； $M_t^{(2)}$ 为 t 时刻的二次平均移动数； N 为计算移动平均数所选的数据个数。

一次移动平均序列总是落后于实际数据序列，出现了滞后偏差；二次移动平均序列也与一次移动平均序列形成滞后偏差。二次移动平均法正是利用这种滞后偏差规律建立线性预测模型，其线性模型为

$$\hat{y}_{t+T} = a_t + b_t T \quad (5-8)$$

式中： t 为目前的序列号； T 为由目前时刻 t 到预测时刻的间隔周期数； \hat{y}_{t+T} 为 $t+T$ 时刻的预测值； a_t 为线性模型的截距； b_t 为线性模型的斜率。 a_t 、 b_t 的可由以下公式求得：

$$a_t = 2aM_t^{(1)} - M_t^{(2)} \quad (5-9)$$

$$b_t = \frac{2}{N-1}(M_t^{(1)} - M_t^{(2)}) \quad (5-10)$$

5.2.3 指数平滑预测法

加权移动平均法中，若加的权值几何减小，就演变成指数平滑预测法。指数平滑是通过某种平均方式，消除历史统计序列中的随机波动，找出其主要发展趋势。指数平滑法有一次指数平滑法（简单指数平滑法）、二次指数平滑法、三次

指数平滑法和高次指数平滑法。高次指数平滑一般很少使用。

指数平滑法认为数据的重要程度按时间上的远近呈非线性递减,其基本原则是越近期的数据,对未来的影响就越大。它通过引入平滑系数 β ,对每一期的消耗量进行平滑,来预测下一期的消耗量。指数平滑法另一个原则是,不断用预测误差来纠正新的预测值,即“误差反馈”原理。

1. 一次指数平滑法

设 y_i ($i=1, \dots, t$) 为第 i 期消耗的实际值; \hat{y}_i ($i=1, \dots, t$) 为第 i 期消耗的预测值; 时间序列为 y_1, y_2, \dots, y_t 。

一次指数平滑法公式表示如下:

$$S_t^{(1)} = \beta y_t + (1 - \beta) S_{t-1}^{(1)} \quad (5-11)$$

或

$$S_t^{(1)} = S_{t-1}^{(1)} + \beta(y_t - S_{t-1}^{(1)}) \quad (5-12)$$

式中: $S_t^{(1)}$ 为一次平滑值; β 为平滑系数, 且 $0 < \beta < 1$ 。

由式(5-11)可以看出,只要知道当期的实际值和上一期的指数平滑值,就可用 β 和 $1 - \beta$ 加权求和,得出当期的指数平滑值。由此可见,利用指数平滑法不需要很多的时间序列数据,也不需要确定几个权重,只要寻找 β 值即可。

与移动平均预测法相似,对于平滑系数 β 的选择同样重要,其值越大反映越灵敏,越易受随机因素的干扰,预测风险就越大。由式(5-11)可以看出, β 的大小规定了在新预测值中新数据和原预测值所占的比例。 β 越大,新数据所占的比重就越大,原预测值所占的比例就越小;反之,则占比例就越大。从式(5-12)可以看出,新预测值是根据预测误差对原预测值进行修正得到的。 β 的大小则体现了修正的幅度, β 值越大,修正幅度越大; β 值越小,修正幅度越小。

因此, β 值既代表预测模型对时间序列数据变化的反应速度,又决定了预测模型修正误差的能力。

若选取 $\beta = 0$, 则 $S_t^{(1)} = S_{t-1}^{(1)}$ 。即下期预测值就等于本期预测值,在预测过程中不考虑任何新信息。

若选取 $\beta = 1$, 则 $S_t^{(1)} = y_{t-1}$ 。即下期预测值等于本期观测值,完全不相信过去的信息,这两种极端情况很难做出正确的预测。因此, β 值应根据时间序列的具体发生情况在 $0 \sim 1$ 范围内选择。选择时遵循下列原则:

(1) 如果时间序列波动不大,比较平衡,则 β 应取小一些,如 $0.1 \sim 0.3$,以减少修正幅度,使预测模型能包含较长时间序列的信息。

(2) 如果时间序列具有迅速且幅度较大的波动, 则 β 应大一些, 如 $0.6 \sim 0.8$, 使预测模型灵敏度高一些, 以便迅速跟上数据的变化。

若近期数据作用越大, 则 β 值应取大些; 反之, β 值应取小些。在实用时, 多取几个 β 值进行试算, 看哪一个预测误差较小, 就采用哪一个 β 值作为权重。

接下来介绍平滑系数 β 和初始预测值的确定方法。

2. 平滑系数的确定

1) 直观法

直观法是根据序列的变化速度来确定平滑系数, 又称为主观法。对序列的实际变化进行观察, 若序列变化缓慢, 则宜选择减小的平滑系数; 反之, 宜选择接近 1 的平滑系数。

2) 模拟法

模拟法又称为客观法, 对于一个确定的平滑系数 β 和约定俗成的初始预测值 $S_0^{(1)}$, 计算出 $S_t^{(1)}$ 的序列预测值, 从而可以计算出预测误差:

$$e_t^{(1)} = y_t - S_{t-1}^{(1)} \quad (5-13)$$

令

$$Q(\beta) = \sum_{t=1}^n [e_t^{(1)}]^2 = \sum_{t=1}^n [y_t - S_{t-1}^{(1)}]^2 \quad (5-14)$$

从而通过求出 $Q(\beta)$ 的最小值来确定平滑系数 β 。通常情况下, $Q(\beta)$ 在 $(0,1)$ 区间内是单峰函数时, 可利用黄金分割点法来计算, 其步骤如下:

(1) 根据 β 的取值范围 $0 < \beta < 1$ 和黄金分割法原则, 第一个 β 记为 β_1 , $\beta_1 = (1-0) \times 0.618 = 0.618$, 然后根据式 (5-14) 计算 $Q(\beta_1)$ 的值。

(2) 选取 $\beta_2 = 1 - 0.618 = 0.382$, 计算 $Q(\beta_2)$ 。

(3) 对 $Q(\beta_1)$ 与 $Q(\beta_2)$ 进行比较。若 $Q(\beta_1) > Q(\beta_2)$, 则最优 β 必在区间 $(0, \beta_1)$, 略去区间 $(\beta_1, 1)$; 反之, 最优 β 必在区间 $(\beta_2, 1)$, 略去 $(0, \beta_2)$ 。

(4) 假定最优 β 在区间 $(0, \beta_1)$, 即 $(0, 0.618)$ 内, 选取与 β_2 对称的点 β_3 ($\beta_3 = 0.618 - 0.382 = 0.236$), 再次计算 $Q(\beta_3)$, 比较 $Q(\beta_2)$ 和 $Q(\beta_3)$ 的值, 再次对区间进行压缩。

(5) 重复以上步骤, 直到得到满意的结果, 从而确定最优 β 。

3. 初始预测值的确定

初始值实质上应为时间序列起点 $t=0$ 以前所有历史数据的加权平均值。在

实际计算中, 初始值确定仅是最初一次。利用式 (5-13) 和式 (5-14) 进行外推预测, 还必须确定初始预测值 $S_0^{(1)}$ 。可以用以下五种方法确定:

(1) 最简单的方法是取前几个数的平均值作为初始值。一般取 3~5 个数的算术平均值。

(2) 若样本序列 $y_1, y_2, y_3, \dots, y_n$ 较长, 则可取 $S_0^{(1)} = \sum_{j=1}^n y_j / n$ 。

(3) 若样本序列不长也不短, 可取 $S_0^{(1)} = y_1$ 。

(4) 取 y_1 的反向预测值, 运用反向平滑预测, 当 t 变至 1 时, 即可记为初始值 $S_0^{(1)}$ 。

(5) 当序列长度充分大, 初始预测值可以任意确定。

4. 二次指数平滑法

与二次移动平均法相似, 对于有线性趋势序列的外推预测, 也可以利用二次指数平滑法。二次指数平滑法是对一次指数平滑后的序列数据再做一次指数平滑, 其平滑公式为

$$S_t^{(2)} = \beta S_t^{(1)} + (1 - \beta) S_{t-1}^{(2)} \quad (5-15)$$

式中: $S_t^{(2)}$ 为 t 时刻的二次指数平滑值; $S_t^{(1)}$ 为 t 时刻的一次指数平滑值; $S_{t-1}^{(2)}$ 为 $t-1$ 时刻的二次指数平滑值; β 为平滑系数。

同样仿照二次移动平均法, 根据滞后偏差的演变规律建立线性预测模型:

$$\hat{y}_{t+T} = a_t + b_t T \quad (5-16)$$

式中: T 为由目前时刻 t 到预测时刻的间隔周期数; \hat{y}_{t+T} 为 $t+T$ 时刻的预测值; a_t 为线性模型的截距; b_t 为线性模型的斜率。 a_t 、 b_t 的计算公式为

$$a_t = 2S_t^{(1)} - S_t^{(2)} \quad (5-17)$$

$$b_t = \frac{\beta}{1 - \beta} (S_t^{(1)} - S_t^{(2)}) \quad (5-18)$$

5. 指数平滑法预测工作流程

指数平滑法预测工作流程如图 5-1 所示。

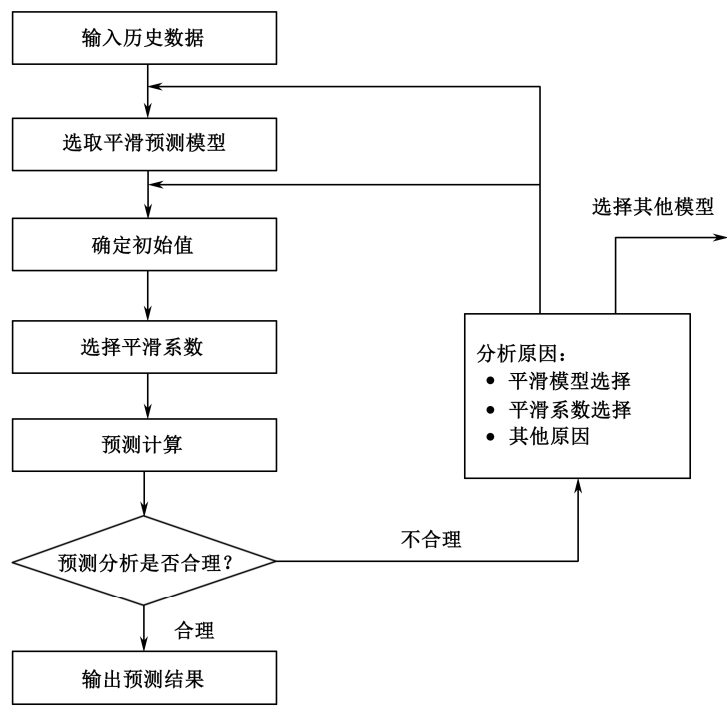


图 5-1 指数平滑法预测工作流程

5.2.4 示例分析

1. 移动平均法

1) 简单移动平均法

设时序为 X_1, X_2, \dots, X_n ，对其中连续 $N(\leq n)$ 个数据点进行算术平均，得 t 时点的移动平均值，记为 M_t ，有

$$M_t = \frac{x_t + x_{t-1} + \dots + x_{t-(N-1)}}{N}$$

采用移动平均值作为预测值 \hat{x}_{t+1} ，则有

$$\hat{x}_{t+1} = M_t = M_{t-1} + \frac{x_t - x_{t-N}}{N} = \hat{x}_t + \frac{x_t - x_{t-N}}{N}$$

由统计数据可知，前 10 个时间段内陀螺组合的需求量如表 5-1 所列，曲线图如图 5-2 所示。

表 5-1 陀螺组合需求统计数据

时间段	1	2	3	4	5	6	7	8	9	10
备件消耗量 y_t	14	13	15	16	15	17	16	15	17	16

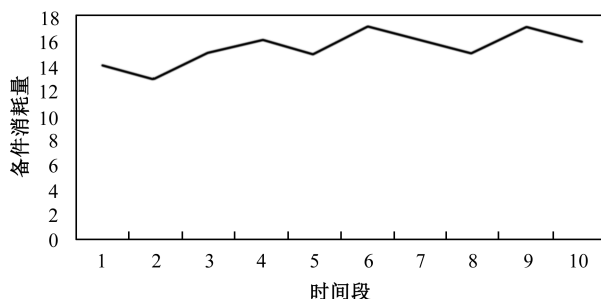


图 5-2 前 10 个时间段备件消耗

通过计算容易得到，当观测点数取 3 时，预测结果误差最小。则由此可预算出第 11 时间段陀螺组合的消耗量 $\hat{x}_{11} = 16$ 个。

预测得到第 11 个时间段的陀螺组合消耗数量为 16 个。

简单移动平均法只适合做近期预测，而且是预测目标的发展趋势变化不大的情况。如果目标的发展趋势存在其他的变化，采用简单移动平均法就会产生较大的预测偏差和滞后。

2) 加权移动平均法

在简单移动平均公式中，每期数据在求平均时的作用是等同的，但是，每期数据所包含的信息量不一样，近期数据包含着更多关于未来情况的信息。因此，把各期数据等同看待是不尽合理的，应考虑各期数据的重要性，对近期数据给予较大的权重，这就是加权移动平均法的基本思想。

设时间序列为 $y_1, y_2, \dots, y_t, \dots$ ，加权移动平均公式为

$$M_{nw} = \frac{w_1 y_t + w_2 y_{t-1} + \dots + w_N y_{t-N+1}}{w_1 + w_2 + \dots + w_N} \quad (t \geq N)$$

利用加权移动平均数来做预测，其预测公式为 $\hat{y}_{t+1} = M_{nw}$ ，即以第 t 期加权移动平均数作为第 $t+1$ 期的预测值。

取 $w_1 = 3, w_2 = 2, w_3 = 1$ ，则易得加权平均法实际消耗值、预测值及相对误差如表 5-2 所列。按预测公式

$$\hat{y}_{t+1} = \frac{3y_t + 2y_{t-1} + y_{t-2}}{3 + 2 + 1}$$

表 5-2 加权移动平均法实际消耗值、预测值及相对误差

时间段	1	2	3	4	5	6	7	8	9	10
备件实际消耗值	14	13	15	16	15	17	16	15	17	16
加权平均预测值				14.17	15.16	15.33	16.17	16.17	15.67	16.17
相对误差/%				11.4	1.0	9.8	1.1	7.8	7.8	1.1

计算得第 11 时间段内陀螺组合的消耗量为

$$\hat{y}_{11} = \frac{3 \times 16 + 2 \times 17 + 15}{3 + 2 + 1} = 16.17 \text{ (个)}$$

移动平均法方法简单，但只做一期预测。一般只对发展变化较平坦，增长趋势不明显，且与以往时期状况关联性不大的时序有效。

2. 指数平滑法

1) 一次指数平滑法

对于一次指数平滑法，有

$$\hat{x}_{t+1} = S_t^{(1)} = \alpha x_t + (1 - \alpha) S_{t-1}^{(1)}$$

式中： α 为平滑系数。

平滑系数描述对过程变化的反应速度， α 越大（接近 1），表示重视近期数据的作用，对过程变化反应越快；描述预测系统对随机误差的修匀能力； α 越小（接近 0），表示重视离现时更远的历史数据的作用，修匀能力越强，但对过程变化的反应越迟钝。

$$\hat{x}_{t+1} = a_0 x_t + a_1 x_{t-1} + a_2 x_{t-2} + \cdots + a_{t-1} x_1 \quad (0 \leq a_i \leq 1, \sum_{i=0}^{t-1} a_i = 1)$$

$$a_0 = \alpha, a_j = \alpha(1 - \alpha)^j, (j = 1, 2, \cdots, t - 1; 0 \leq \alpha \leq 1)$$

$$\hat{x}_{t+1} = \alpha x_t + (1 - \alpha) \hat{x}_t$$

式中： S_t^1 、 S_t^2 分别为 t 时刻的一次指数平滑值和二次指数平滑值。

陀螺组合一次平滑值与二次指数平滑值对比如图 5-3 所示。

陀螺组合

对于陀螺组合，指数平滑值有

$$S_1^{(1)} = x_1 = 14, \quad \alpha = 0.3$$

由 $S_t^{(1)} = \alpha x_t + (1 - \alpha) S_{t-1}^{(1)}$ 可得

$$S_1^{(1)} = 0.3 \times 14 + 0.7 \times 14 = 14$$

$$S_2^{(1)} = 0.3 \times 13 + 0.7 \times 14 = 13.7$$

$$S_3^{(1)} = 0.3 \times 15 + 0.7 \times 13.7 = 14.09$$

$$S_{10}^{(1)} = 0.3 \times 16 + 0.7 \times 15.89 = 15.923$$

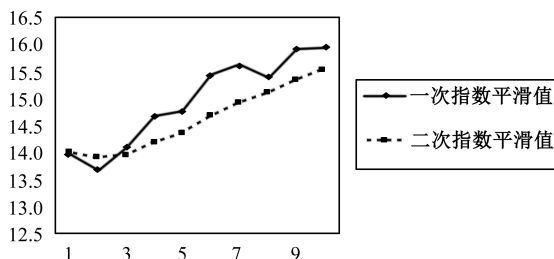


图 5-3 指数平滑法预测结果对比

所以, 第 11 时间段内陀螺组合的消耗预测方程为

$$\hat{x}_{11} = \alpha x_{10} + (1 - \alpha) \hat{x}_{10}$$

代入数据可预测得到第 11 时间段内陀螺组合的消耗量 $\hat{x}_{11} = 15.95$ 个。

2) 二次指数平滑法

用与一次指数求解相似的方法, 可得到陀螺组合在第 10 时间段的二次指数平滑值 $S_{10}^{(2)} = 15.51$ 。一次指数平滑法虽然克服了移动平均法的缺点, 但当时间序列的变动出现直线趋势时, 用一次指数平滑法进行预测仍存在明显的滞后偏差, 也必须加以修正。修正的方法与趋势移动平均法相同, 即再做二次指数平滑, 利用滞后偏差的规律建立直线趋势模型。这就是二次指数平滑法。其计算公式为

$$S_t^{(1)} = \alpha y_t + (1 - \alpha) y_{t-1}^{(1)}$$

$$S_t^{(2)} = \alpha S_t^{(1)} + (1 - \alpha) S_{t-1}^{(2)}$$

式中: $S_t^{(1)}$ 为一次指数的平滑值; $S_t^{(2)}$ 为二次指数平滑值。

当时间序列 $\{y_t\}$ 从某时期开始具有直线趋势时, 类似趋势移动平均法, 可用直线趋势模型, 即

$$\hat{x}_{t+1} = a_{10} + b_{10}T$$

$$a_{10} = 2S_{10}^{(1)} - S_{10}^{(2)}$$

$$b_{10} = \frac{\alpha}{1 - \alpha} [S_{10}^{(1)} - S_{10}^{(2)}]$$

代入数据可得

$$a_{10} = 2S_{10}^{(1)} - S_{10}^{(2)} = 15.09$$

$$b_{10} = \frac{\alpha}{1-\alpha} [S_{10}^{(1)} - S_{10}^{(2)}] = 0.177$$

则预测方程为

$$\hat{x}_{t+1} = a_{10} + b_{10}T = 15.09 + 0.177T$$

代入数据可得 $\hat{x}_{t+1} = 16.86$ 个。

5.3 基于回归分析的备件需求预测方法

由于备件消耗受部件自身可靠性、工作的综合环境（自然环境与诱发环境）、部件使用强度等因素的影响，因此，寻求主要影响因素与备件消耗量之间的关系已成为备件保障需要解决的重要问题之一。回归分析法是从一组观测数据出发，来确定随机变量的定量函数关系，即回归模型。

5.3.1 线性回归概述

在客观世界中普遍存在变量之间的关系，变量之间的关系一般来说可分为确定性的与非确定性的两种。确定性关系是指变量之间的关系可以用函数关系来表达，而非确定性的关系只是一种变量之间的相关关系。回归分析法是一种从事物变化的因果关系出发进行预测的方法，是研究相关关系的一种数学工具，是维修器材消耗预测的常用方法之一。它利用数理统计基本原理，在大量统计数据基础上，通过数据变化规律来推测、判断和描述事物未来的发展趋势。它寻求消耗量与一个或多个影响消耗量变化的自变量之间的关系，依据多组观察数据，找出拟合这些数据点的最佳拟合曲线，依此确定影响消耗量变化的诸因素与消耗量的关系式，从而建立消耗函数模型。

回归分析法是一种应用广泛、理论性较强的定量预测方法。它的基本思路是分析预测对象与有关因素的相互联系，用适当的回归预测模型（回归方程）表达出来，然后根据数学模型预测其未来状态。由于有较严密的理论基础和较成熟的计算分析方法，如果模型建立得当，则可得到比较精确的预测结果。回归分析法的优点如下：

(1)能研究预测对象与相关因素的相互关系，抓住预测对象变化的实质原因，因而预测结果比较可信。

(2)能给出预测结果的置信区间和置信度，从而使预测结果更加完整和客观。

(3) 考虑了相关性, 能运用有关的数理统计方法对回归方程进行统计检验, 从而对预测对象变化的转折点具有一定的鉴别能力。

回归分析法的缺点如下:

(1) 对实际数据一视同仁, 认为各数据对预测对象的影响程度相同, 这是不符合实际的。

(2) 计算工作量较大, 出现新数据时, 一般要重新估计回归方程和进行相关分析。

(3) 选取回归变量时的主次要因素在实际建模时较难把握, 变量因素的量化也是一个难点。

(4) 回归分析法往往只能考虑少数几种主要影响因素, 而略去了许多未考虑的因素, 影响了预测的准确性。

5.3.2 一元线性回归模型

一元线性回归预测是指根据成对的两个变量数据分布大体上呈直线趋势时, 采用适当的计算方法, 找到两者之间特定的经验公式, 即一元线性回归模型, 然后根据自变量的变化, 来预测因变量发展变化的方法。

1. 模型建立

设 x 是可控变量, Y 是依赖于 x 的随机变量, 它们之间的关系为

$$Y = \alpha + \beta x + \varepsilon \quad (5-19)$$

式中: α 、 β 为常数; ε 服从正态分布 $N(0, \sigma_2)$ 。

x 与 Y 为这种关系称为一元线性回归 (模型)。当 x 取固定值时, Y 服从正态分布 $N(\alpha + \beta x, \sigma_2)$ 。上式两边取数学期望得

$$E(Y) = \alpha + \beta x \quad (5-20)$$

式中: α 、 β 为回归系数。

式 (5-20) 为 Y 对 x 的回归直线方程。

在实际实验中, 对变量 x 和 Y 做 n 次试验观察, 并且假定在 x 的各个值上对 Y 的观察是相互独立的, 得到 n 对试验值。在平面笛卡儿坐标系中, 画出坐标为 (x_i, y_i) ($i=1, 2, \dots, n$) 的 n 个点, 它们所构成的图像称为点图。如果 n 很大时, 点图中的 n 个点分布在一条直线附近, 直观上可以认为 x 与 Y 的关系具有式 (5-20) 的模型。 Y 相应于 x_1, x_2, \dots, x_n 的观察值 y_1, y_2, \dots, y_n 可看成 Y_1, Y_2, \dots, Y_n 的试验值, 有

$$Y_i = \alpha + \beta x_i + \varepsilon_i \quad (i=1, 2, \dots, n) \quad (5-21)$$

式中: ε_i 服从正态分布 $N(0, \sigma_2)$, 且 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 相互独立。

式 (5-24) 通常称为线性模型。显然, Y 服从正态分布 $N(\alpha + \beta x, \sigma_2)$, 且 y_1, y_2, \dots, y_n 相互独立。 Y_n 的独立性意味着在 x 的各个值上对 y 的观察是相互独立的。

在一元线性回归中主要解决下列一些问题:

- (1) 用试验值 $(x_i, Y_i) (i=1, 2, \dots, n)$ 对未知参数 α 、 β 和 σ_2 做点估计;
- (2) 进行检验;
- (3) 在 $x = x_0$ 处对 Y 做预测, 备件的需求预测用的就是这种方法。

2. 参数估计

已知变量 x 、 Y 的行对试验值 $(x_i, Y_i) (i=1, 2, \dots, n)$, 先用最小二乘法求出 α 、 β 的估计值。做离差平方和:

$$Q = \sum_{i=1}^n (y_i - Y_i)^2 = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2 \quad (5-22)$$

选择参数 α , β 使 Q 达到最小, 即

$$Q = \min \left(\sum_{i=1}^n (y_i - \alpha - \beta x_i)^2 \right)$$

经过整理, 回归系数 α 、 β 可以用下式计算:

$$\hat{\beta} = \frac{\sum_{i=1}^s (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^s (x_i - \bar{x})^2} \quad (5-23)$$

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x} \quad (5-24)$$

3. 模型检验

相关系数用于描述变量 x 和 y 的线性相关的程度, 并常用 γ 来表示, γ 的值在 $[-1, 1]$ 范围。 γ 的绝对值越接近于 1, 表示 x 和 y 之间的线性关系越密切。当 $\gamma > 0$ 时, 两者呈正相关; 当 $\gamma < 0$ 时, 两者呈负相关。 γ 的值接近于 0, 两者没有线性相关关系。 γ 的计算公式为

$$\gamma = \frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}} \quad (5-25)$$

式中

$$L_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}), \quad L_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2, \quad L_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2$$

当 $\gamma=1$ 时, 可以准确预测 y 值。此时若 x 不是随机变量, 则 y 也不是随机变量。这种情况在研究中是不多见的。

当 $\gamma=0$ 时, 回归一点作用也没有, 即用 x 的线性函数完全不能预测 y 的变化, 但这时 x 与 y 之间还可能存在着非线性的关系。

当 $0 < \gamma < 1$ 时, 情况介于上述二者之间。 x 的线性函数对预测 y 的变化有一定作用, 但不能准确预测。这说明, y 还受其他一些因素, 包括随机误差的影响。

4. 置信区间求取

在实际问题中, $|\gamma|=1$ 的情况几乎是没有的, 预测值与实际实现值总会有偏差, 预测的愿望总希望能尽可能接近实际值, 不超出一定的范围或区间。在统计学中, 要求实际值位于这个区间范围的概率达到 95% 以上, 这个区间即预测值的置信区间。置信区间说明回归模型的适用范围或精确程度, 一般地, 在统计学的误差分析中, 数据点在回归直线附近大致接近正态分布时, 规定这个区间为 $y \pm 2\sigma$, σ 为标准离差, 其计算公式为

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (5-26)$$

这样就可以得出置信区间上、下限的两条控制线, 即

$$y_1 = \alpha + \beta x \pm 2\sigma \quad (5-27)$$

5.3.3 多元线性回归模型

多元回归分析方法是通过建立统计模型研究多变量间相互关系的密切程度、结构状态、模型预测的一种有效的工具。多元线性回归预测是一元线性回归预测的直接推广, 其包含一个因变量和两个或两个以上的自变量。

多元线性回归分析的基本思想: 多元回归分析是处理变量之间的一种统计方法和技术, 它研究的主要对象是客观事物间的统计关系, 它是建立在对客观事物进行大量试验和观察的基础上的, 用来寻找隐藏在那些看上去不确定现象中的统计规律性的统计方法。

1. 模型建立

设因变量 y 与自变量 x_1, x_2, \dots, x_m 共有 n 组实际观测数据, 分别为

$$(x_{i1}, x_{i2}, \dots, x_{im}; y_i) \quad (i=1, 2, \dots, n)$$

y 是一个可观测的随机变量, 它受到 m 个非随机因素 x_1, x_2, \dots, x_m 和 ε 随机因素的影响。若 y 与 x_1, x_2, \dots, x_m 有线性关系

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon \quad (5-28)$$

即

$$\begin{cases} y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_m x_{1m} + \varepsilon_1 \\ y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_m x_{2m} + \varepsilon_2 \\ \vdots \\ y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_m x_{nm} + \varepsilon_n \end{cases}$$

式中: y 为因变量; x_1, x_2, \dots, x_m 为自变量; $\beta_0, \beta_1, \dots, \beta_m$ 是未知参数; ε 是均值为零、方差大于 0 的不可观测的随机变量, 称为误差项, 并通常假定 $\varepsilon \sim N(0, \sigma^2)$ 。

2. 参数估计

采用最小二乘法估计总体参数 $\beta = (\beta_0, \beta_1, \dots, \beta_m)^T$, 其估计量 $B = (b_0, b_1, \dots, b_m)^T$, 总体参数的最小二乘估计量为

$$B_{LS} = (X^T X)^{-1} X^T Y \quad (5-29)$$

3. 模型检验

1) 拟合优度检验 (R 检验)

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (5-30)$$

根据回归模型的自由度 $n-m$ 和给定的显著性水平 α 值, 查相关检验表进行判别。如果 $R \geq R_\alpha(n-m)$, 则预测可靠; 如果 $R < R_\alpha(n-m)$, 则预测结果不能使用。

2) 回归方程的显著性检验 (F 检验)

$$F = \frac{\sum (y_i - \bar{y})^2 / m}{\sum (y_i - \hat{y}_i)^2 / n - m - 1} \quad (5-31)$$

给定显著水平 α , 查 F 分布表得临界值 $F_\alpha(m, n-m-1)$, 计算 F 的观测值 F_0 。若 $F_0 \geq F_\alpha(m, n-m-1)$, 则存在显著的线性关系; 若 $F_0 < F_\alpha(m, n-m-1)$, 则认

为线性关系不显著。

3) 回归系数的显著性检验 (t 检验)

$$t_j = \frac{b_j}{\sqrt{c_{jj} \sum (y_i - \hat{y}_i)^2 / (n - m - 1)}} \quad (j = 1, 2, \dots, m) \quad (5-32)$$

式中: c_{jj} 为矩阵 $(\mathbf{X}\mathbf{X})^{-1}$ 主对角线上的第 j 个元素。

给定显著水平 α , 查 t 分布表得临界值 $t_{\frac{\alpha}{2}}(n - m - 1)$, 计算 t_j 的统计量 $|t_j|$, 当 $|t_j| > t_{\frac{\alpha}{2}}(n - m - 1)$, 存在显著的线性关系; 否则, 认为线性关系不显著。

5.3.4 示例分析

CC1-63V-0.01 μ F 瓷介电容器是某型鱼雷配备较多的部件。经分析, 该补部件消耗量主要与装备通电工作时间的长短有关, 还与一些辅助因素 (如周围环境条件、工作人员的责任心、保管质量等) 有关。已知 300 枚该型装备 2000—2009 年的保障数据如表 5-3 所列。若预计该装备 2010 年的通电工作时间为 120h, 试预测 2010 年该备件的消耗量。

表 5-3 某型鱼雷装备 2000—2009 年保障数据

年份	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
通电工作时间/h	105	99	102	83	52	67	79	97	100	93
器材消耗数量/个	15	13	14	10	6	8	11	13	14	12

根据表 5-3, 进行统计计算可得

$$\sum_{i=1}^{10} x_i = 877, \sum_{i=1}^{10} x_i^2 = 79611, \sum_{i=1}^{10} x_i \cdot y_i = 10614, \sum_{i=1}^{10} y_i = 116, \sum_{i=1}^{10} y_i^2 = 1420, n = 10$$

将这些数据代入式 (5-26) 和式 (5-27) 得到结果: $\alpha = -2.75, \beta = 0.1634$ 。

由此得到回归模型为

$$y = -2.75 + 0.1634x$$

则备件消耗量为

$$y = -2.75 + 0.1634 \times 120 \approx 17 \text{ 个}$$

采用相关性检验法对回归进行显著性检验。根据式 (5-28) 计算相关系数, 可得 $\gamma \approx 0.98$ 。根据保障实际情况, 取置信水平 $\alpha = 0.001$, $n - 2 = 8$, $\gamma = 0.98$, 查相关系数, 可得

$$\gamma_{8,0.001} = 0.8721$$

由于 $0.98 > 0.8721$ ，因此认为回归效果是显著的。

说明备件消耗数量与通电工作时间存在很强的相关关系，即该备件的消耗数量 98% 取决于工作运行时间的长短，而只有 2% 取决于其他原因。

应用上述回归模型进行预测计算，结果如表 5-4 所列。

表 5-4 瓷介电容器备件的预测数据

年份	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
预测消耗数量/个	14.4	13.4	13.9	10.8	5.7	8.2	10.2	13.1	13.6	12.4

根据式 (5-29) 求得 $\sigma \approx 0.48$ ，置信区间为

$$y = -2.75 + 0.1634x \pm 0.96$$

因此，上述维修器材 2010 年的预测结果为 17 ± 1 个。

由此可以看出，在相关系数为 98% 时，其预测误差是比较小的。

回归分析法能给出预测结果的置信区间和置信度，从而使预测结果更加完整和客观，但需要有大量数据，且对数据没有区别的对待，量化困难，运算量大。另外，回归分析法所得到的回归预测方程往往只能考虑少数几种主要影响因素，而略去了许多未考虑的因素，影响了预测的准确性。回归分析法主要适用于在很长时期内备件需求量呈平稳分布的情况。

第 6 章

基于灰色模型的武器装备备件 需求预测方法



武器装备备件需求是由装备战备任务、储存条件、保障情况、天气、环境、经济等因素共同影响、制约的结果，如此众多的因素很难用语言和指标表达清楚，其对备件需求的潜在而复杂的影响更是无法精确计算的，这反映备件需求具有明显的灰色性。

6.1 灰色预测方法概述

6.1.1 灰色系统简介

灰色系统理论是我国著名学者邓聚龙教授于 1982 年创立的新兴横断学科中提到的，是研究灰色系统建模、预测、分析、控制和决策的理论。它的研究对象是“部分信息已知，部分信息未知”的“小样本”“贫信息”不确定性系统，通过对“部分”已知信息的生成、开发去了解、认识现实世界，实现对系统运行行为和演化规律的正确把握和描述。灰色系统模型对试验观测数据及其分布没有特殊的要求和限制，构造简易，适用性强，应用范围广。

灰色系统理论认为：尽管客观系统表象极为复杂，表征数据看似杂乱无章，但是系统潜藏着内在规律，因素具有整体功能。灰色系统理论是从系统的角度出发，按某种逻辑推理和理性认识来研究信息间关系。它认为客观世界是信息的世界，既有大量已知信息，又有不少未知信息或非确知信息。用“黑”表示信息未知，用“白”表示信息完全明确，用“灰”表示部分信息明确、部分信息不明确，“信息不完全”是“灰”的基本含义。相应地，信息未知的、非确知的系统称为黑色系统，信息完全明确的系统称为白色系统，部分信息明确、部分信息不明确的系统或既含有已知信息又含有未知信息的系统称为灰色系统。

概率统计、模糊数学和灰色系统理论是三种常用的不确定性系统的研究方法，研究对象都具有某种不确定性。灰色系统理论着重研究“外延明确，内涵不明确”的对象，着重研究概率统计、模糊数学所不能解决的“小样本、贫信息不确定”问题。灰色系统理论利用灰色信息建立微分方程并构造灰色模型以达到对未知领域的探索，它能揭示系统内部连续发展的变化过程。它避开了讨论系统模型参数之间的关系，从其本身的内在规律着手，通过已知信息来预测未知领域，从而达到了解整个系统的目的。它依据信息覆盖，利用那些有限的、不完全确切的、表示系统行为特征的原始数据序列做生成变换后，建立微分方程，推求或预测系统特征的全貌或发展趋势，寻求现实规律。其最大特点是“少数据建模”，对样本数量没有特殊要求，允许数据为非经典分布。该理论建模方法简单，意义明确，所需信息量少，运算简捷方便，预测精度较高。

6.1.2 灰色系统理论基础

1. 灰数

灰数是灰色系统理论中最基本的概念，只知道大概范围而不知道其确切值的数称为灰数，通常用符号“ \otimes ”表示。

根据不确切程度，灰数可分为仅有下界的灰数、仅有上界的灰数和区间灰数等。当 $\otimes \in [-\infty, +\infty]$ ， \otimes 称为黑数；当 $\otimes \in [a, \bar{a}]$ 且 $\underline{a} = \bar{a}$ 时， \otimes 称为白数。

灰数又可分为本征灰数与非本征灰数。前者是指不能或暂时还不能找到一个白数作为其“代表”的灰数；而后者是指凭先验信息或某种手段，可以找到一个白数作为其代表的灰数，称此白数为相应灰数的白化值，记为 $\tilde{\otimes}$ ，并用 $\otimes(a)$ 表示以 a 为白化值的灰数。

2. 累加生成算子

累加生成算子在灰色系统理论中占有极其重要的地位,利用该算子可使灰色过程由灰变白,得出灰量积累过程的发展态势,从凌乱的数据中得出所蕴含的积分特性或规律。

定义 6.1 设 $X^{(0)}$ 为原始序列, $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, D 为序列算子, 则有 $X^{(0)}D = (x^{(0)}(1)d, x^{(0)}(2)d, \dots, x^{(0)}(n)d)$ 式中

$$x^{(0)}(k)d = \sum_{i=1}^k x^{(0)}(i) \quad (k=1, 2, \dots, n)$$

则称 D 为 $X^{(0)}$ 的一次累加生成算子, 记为 1-AGO。

3. 累减生成算子

累减生成算子是累加生成算子的逆运算,是对累加生成的还原,经常用于获取增量信息。

定义 6.2 设 $X^{(0)}$ 为原始序列, $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, D 为序列算子, 则有 $X^{(0)}D = (x^{(0)}(1)d, x^{(0)}(2)d, \dots, x^{(0)}(n)d)$ 式中

$$x^{(0)}(k)d = x^{(0)}(k) - x^{(0)}(k-1) \quad (k=2, 3, \dots, n)$$

则称 D 为 $X^{(0)}$ 的一次累加生成算子。

4. 序列的光滑性

灰色系统理论所研究的序列是由离散的单个点构成的,通常意义下无导数可言,因此,不能像用导数研究光滑连续函数一样研究序列的光滑性,而是通过引入光滑比的概念来研究序列的光滑性。

定义 6.3 称 $\rho(k) = \frac{x(k)}{\sum_{i=1}^{k-1} x(i)}$ ($k=2, 3, \dots, n$) 为序列 X 的光滑比。

光滑比用序列中第 k 个数据 $x(k)$ 与前 $k-1$ 个数据之和来考查序列 X 中数据变化是否平稳。

定义 6.4 若序列 X 满足

$$\begin{aligned} \frac{\rho(k+1)}{\rho(k)} &< 1 \quad (k=2, 3, \dots, n-1) \\ \rho(k) &\in [0, \varepsilon] \quad (k=3, 4, \dots, n) \end{aligned}$$

$$\varepsilon < 0.5$$

则称 X 为准光滑序列。

5. 累加生成的灰指数律

定义 6.5 设序列 $X = (x(1), x(2), \dots, x(n))$:

如果 $\forall k, \sigma(k) = \frac{x(k)}{x(k-1)} \in (0, 1]$, 则称序列 X 具有负的灰指数规律。

如果 $\forall k, \sigma(k) = \frac{x(k)}{x(k-1)} \in (1, b]$, 则称序列 X 具有正的灰指数规律。

如果 $\forall k, \sigma(k) = \frac{x(k)}{x(k-1)} \in [a, b], (b-a = \delta)$, 则称序列 X 具有绝对灰度为 δ 的

灰指数规律。

当 $\delta < 0.5$ 时, 则 X 具有准指数规律。

定理 6.1 设序列 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ 为非负准光滑序列, $X^{(0)}$ 的一次累加生成序列 $X^{(1)}$ 具有准指数规律。在使用各种灰色预测模型对序列进行预测时, 首先必须验证该序列是否为非负准光滑序列。若不能满足这个条件, 则需对原始数列进行预处理。

6.1.3 灰色预测方法

灰色预测方法是 20 世纪 90 年代后期的一种新型非线性预测技术。灰色预测的基本思想是将实际产生的数据视为在一定幅值范围、一定时区内变化的灰色量, 将其变化过程看成是一定幅区、一定时区内变化的灰色过程; 同时, 灰色系统理论认为这些被表象所迷惑、所掩盖的离乱的灰色量对广义的能量系统而言, 具有某种整体功能, 是有序的, 其中必然隐藏着某种内在的规律, 通过对原始数据的处理, 便可得到依据内在规律而产生的结果。灰色预测法具有不需要确定数据变动是否服从某一类型分布, 不需要大的样本统计量就可以进行预测的优势。

灰色预测模型不同于数理统计方法, 不需要确定数据变动是否服从正态分布, 不需要大的样本统计量, 不需要根据数据变化随时改变预测模型, 而是通过累加、累减或者级比生成技术直接将时间序列转化为微分方程。由于微分方程模型具有比代数模型、离散模型更高的光滑度, 根据预测的连续性原理, 理论上它具有更高的预测精度。灰色预测方法用于处理研究对象信息量不充分而出现的不确定性(灰色性)的问题, 适用于解决小样本、贫信息和不确定性的问题, 即少

数据建模。因而该方法对样本量的多少和样本有无规律都同样适用，而且计算量小，十分方便实用。其优势在于短期预测，缺点是对长期预测和波动较大数据序列的拟合较差。

6.2 基于一般序列 GM(1,1)模型的备件需求预测方法

灰色预测模型是灰色系统理论的重要组成部分，灰色模型的一般形式是 GM(n, h)，其中 n 为模型新建立的微分方程的阶数， h 为变量个数。基于灰色理论的 GM(1,1)模型的预测，称为灰色预测，它已广泛应用于经济、农业、气象、社会、生态、水利、生物等领域。

6.2.1 模型建立

灰色预测的基本思想是：首先将随机变量当作在一定范围内变化的灰色量，将随机过程当作在一定范围、一定时间内变化的灰色过程，将原始数据进行灰色生成(AGO)，使其随机因素弱化；然后对生成数列建立白化形式的微分方程，求出方程的解数列；最后进行累减生成还原，得到原始序列的预测值。

GM(1,1)建模的具体步骤如下。

1. 建模可行性分析

(1) 根据所收集到的信息数据，构造一个原始的时间序列：

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)) \quad (x^{(0)}(k) \geq 0; k = 1, 2, \dots, n)$$

(2) 对 $X^{(0)}$ 进行准光滑性检验。

对序列 $X^{(0)}$ ，光滑比为
$$\rho(k) = \frac{x^{(0)}(k)}{\sum_{i=1}^{k-1} x^{(0)}(i)}$$

若有

$$\frac{\rho(k+1)}{\rho(k)} < 1 \quad (\rho(k) \in [0, \varepsilon], \quad \varepsilon < 0.5)$$

则称 $X^{(0)}$ 为准光滑序列。

(3) 做一次累加生成，得 $X^{(0)}$ 的 1-AGO 序列：

$$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \quad (6-1)$$

式中

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) \quad (k=1, 2, \dots, n)$$

(4) 对 $X^{(1)}$ 进行准指数规律检验。

对序列 $X^{(1)}$ ，级比为

$$\sigma(k) = \frac{x^{(1)}(k)}{x^{(1)}(k-1)}$$

若 $\forall k, \sigma(k) \in [a, b]$ ($b-a=\delta, \delta < 0.5$)，则称序列 $X^{(1)}$ 具有准指数规律。

设 $X^{(0)}$ 为非负准光滑序列，则 $X^{(0)}$ 的一次累加生成序列 $X^{(1)}$ 具有准指数规律。这是灰色系统建模的理论基础。也就是说，只有原始序列满足了准光滑性的条件，1-AGO 序列 $X^{(1)}$ 满足准指数规律，才能进行建模。一般情况下，非负序列经过累加生成后可得到准光滑序列，波动性和随机性被明显弱化，确定性增强，呈现出近似的指数增长规律。原始序列越光滑，累加生成后指数规律就越明显。对于满足光滑条件的序列，在生成时间序列的基础上可以建立灰色微分模型。累加生成与均值生成都能提高序列的光滑度，因此有时做完累加生成后还可以再做一次均值生成。

2. 数据变换处理

数据变换处理的原则是经过处理后的原始序列满足准光滑性、1-AGO 序列满足准指数规律，从而可以建立 GM(1,1)模型。对于准光滑性和准指数规律检验不合格的序列，必须做数据处理，使变换后的序列的光滑比和级比满足要求。通常的数据变换有平移变换、对数变换、方根变换、加权变换等。

3. GM(1,1)建模

灰色预测模型 GM(1,1)的建模过程如下：

(1) 假设 $X^{(0)}$ 为非负原始数据序列 $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$ ，其中 $x^{(0)}(i)$ 对应于时刻 i 的系统输出。

(2) 对 $X^{(0)}$ 做一次灰色累加生成操作得到一个新的生成数据序列 $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$ ，其中， $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$ 。

(3) 由新数据序列 $X^{(1)}$ 建立灰色模型 GM(1,1)，对应的白化微分方程为

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b \quad (6-2)$$

灰色微分方程为

$$x^{(0)}(k) + az^{(1)}(k) = b \quad (6-3)$$

式中：\$a\$ 为发展系数；\$b\$ 为灰作用量；\$z^{(1)}(k)\$ 为 \$x^{(1)}(k)\$ 在 \$[k-1, k]\$ 上的背景值，且有

$$z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k-1) + x^{(1)}(k)) \quad (6-4)$$

(4) 运用最小二乘法估计参数 \$a\$ 和 \$b\$，参数 \$\hat{a} = (\hat{a}, \hat{b})^T\$ 的最小二乘参数估计为

$$\hat{a} = (\hat{a}, \hat{b})^T = (B^T B)^{-1} B^T Y \quad (6-5)$$

式中

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}; Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}。$$

(5) 白化微分方程的时间响应函数为

$$\hat{x}^{(1)}(t) = \left[x^{(1)}(1) - \frac{b}{a} \right] e^{-a(t-1)} + \frac{b}{a}$$

离散化的时间响应序列为

$$\hat{x}^{(1)}(k) = \left[x^{(1)}(1) - \frac{b}{a} \right] e^{-a(k-1)} + \frac{b}{a} (k=1, 2, \dots) \quad (6-6)$$

(6) 对得到的建模预测序列 \$\hat{x}^{(1)}\$ 进行一次累减，还原得到原始数据的拟合值为 \$\hat{X}^{(0)} = \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n)\}\$

式中

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1) = (1 - e^a) \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak}$$

其中：\$\hat{x}^{(0)}(1) = x^{(0)}(1)\$。

4. 模型精度检验

残差检验、关联度检验和后验差检验方法都是通过对残差的考察来判断模型的精度的，此处采用基于后验差检验方法来检验模型精度，即对残差分布的统计特性进行检验。

设 $X^{(0)}$ 为原始序列, $\hat{X}^{(0)}$ 为相应的模拟序列, $\varepsilon^{(0)}$ 为残差序列, 则

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x^{(0)}(k), \quad S_1^2 = \frac{1}{n} \sum_{k=1}^n (x^{(0)}(k) - \bar{x})^2 \quad (6-7)$$

分别为 $X^{(0)}$ 的均值、方差。

残差序列

$$\begin{aligned} \varepsilon^{(0)} &= (\varepsilon(1), \varepsilon(2), \dots, \varepsilon(n)) \\ &= \left(x^{(0)}(1) - \hat{x}^{(0)}(1), x^{(0)}(2) - \hat{x}^{(0)}(2), \dots, x^{(0)}(n) - \hat{x}^{(0)}(n) \right) \\ \bar{\varepsilon} &= \frac{1}{n} \sum_{k=1}^n \varepsilon(k), \quad S_2^2 = \frac{1}{n} \sum_{k=1}^n (\varepsilon(k) - \bar{\varepsilon})^2 \end{aligned} \quad (6-8)$$

分别为残差的均值、方差。

$C = \frac{S_2}{S_1}$ 称为均方差比值, 对于给定的 $C_0 > 0$, 当 $C < C_0$ 时, 称模型为均方差比合格模型。

$p = P\left(|\varepsilon(k) - \bar{\varepsilon}| < 0.6745 S_1\right)$ 称为小误差概率, 对于给定的 $p_0 > 0$, 当 $p > p_0$ 时, 称模型为小误差概率合格模型。

C_0 和 p_0 的一组取值确定了检验模型模拟精度的一个等级, 常用的精度等级见表 6-1。

表 6-1 精度等级

精度等级 \ 检验方法	均方差比值 C_0	小误差概率 p_0
一级	0.35	0.95
二级	0.50	0.80
三级	0.65	0.70
四级	0.80	0.60

5. 预测

若精度检验等级为三级或四级, 则需采用残差或模型群等方法进行模型修正; 若为一级或二级, 则可直接用该模型进行数列预测。在装备数量、工作任务量等不变的情况下, 所得预测值可作为下年度备件需求量的参考值。

GM(1,1)模型预测流程如图 6-1 所示。

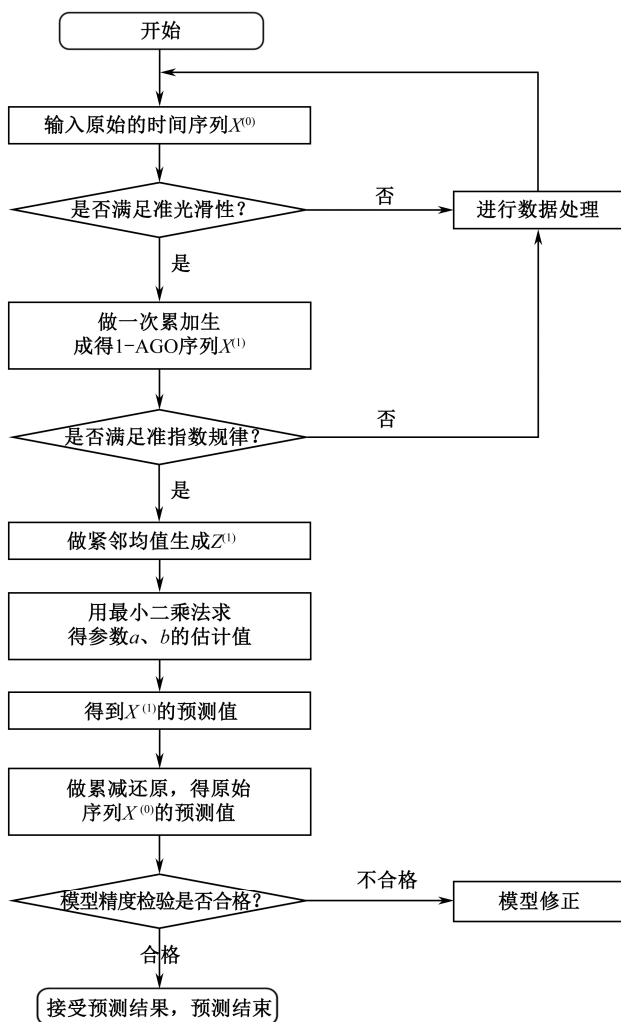


图 6-1 GM(1,1)模型预测流程

6.2.2 示例分析

某型导弹整流罩弹垫 2003—2007 年的消耗量统计数据如表 6-2 所列, 试预测 2008 年该整流罩弹垫的需求量。

表 6-2 某备件 2003—2007 年的消耗数量

年份	2003	2004	2005	2006	2007
备件数量/个	83	81	77	79	85

解：（1）建模准备。

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(5)) = (83, 81, 77, 79, 85)$$

做累加生成得 1-AGO 序列

$$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(5)) = (83, 164, 241, 320, 405)$$

对 $X^{(0)}$ 做准光滑性检验。

由

$$\rho(k) = \frac{x(k)}{\sum_{i=1}^{k-1} x(i)} = \frac{x^{(0)}(k)}{x^{(1)}(k-1)}$$

可得 $\rho(2) \approx 0.98$, $\rho(3) \approx 0.47 < 0.5$, $\rho(4) \approx 0.33 < 0.5$, $\rho(5) \approx 0.27 < 0.5$, 当 $k > 2$ 时, 满足准光滑性条件。

再对 $X^{(1)}$ 做准指数规律检验

由

$$\sigma(k) = \frac{x(k)}{x(k-1)}$$

可得 $\sigma(2) \approx 1.98$, $\sigma(3) \approx 1.47$, $\sigma(4) \approx 1.33$, $\sigma(5) \approx 1.27$, 当 $k > 2$ 时, $\sigma(k) \in [1, 1.5]$, $\delta = 0.5$, 满足准指数规律。

故不需要进行数据变换处理就可对 $X^{(0)}$ 建立 GM(1,1)模型。

（2）建模与检验。

对 $X^{(1)}$ 做紧邻均值生成得序列:

$$Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), z^{(1)}(4)) = (123.5, 202.5, 280.5, 362.5)$$

$$B = \begin{pmatrix} -123.5 & 1 \\ -202.5 & 1 \\ -280.5 & 1 \\ -362.5 & 1 \end{pmatrix}, \quad Y = \begin{pmatrix} 81 \\ 77 \\ 79 \\ 85 \end{pmatrix}, \quad \hat{a} = (B^T B)^{-1} B^T Y = \begin{pmatrix} -0.0178 \\ 76.1808 \end{pmatrix}$$

确定时间响应式:

$$\hat{x}^{(1)}(k+1) = 4362.8202e^{0.0178k} - 4279.8202$$

求 $X^{(1)}$ 的模拟值

$$\begin{aligned}\hat{X}^{(1)} &= (\hat{x}^{(1)}(1), \hat{x}^{(1)}(2), \dots, \hat{x}^{(1)}(5)) \\ &= (83.0000, 161.3571, 241.1238, 322.3255, 404.9879, 489.1374)\end{aligned}$$

还原求出 $X^{(0)}$ 的模拟值

$$\hat{X}^{(0)} = (\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(5)) = (83.0000, 78.3571, 79.7667, 81.2017, 82.6624)$$

最后, 分析误差, 进行精度检验。

残差序列为:

$$\varepsilon^{(0)} = (\varepsilon^{(0)}(1), \varepsilon^{(0)}(2), \dots, \varepsilon^{(0)}(5)) = (0, -2.6429, 2.7667, 2.2017, -2.3376)$$

相对误差序列为:

$$\Delta = \left(\left| \frac{\varepsilon^{(0)}(1)}{x^{(0)}(1)} \right|, \left| \frac{\varepsilon^{(0)}(2)}{x^{(0)}(1)} \right|, \dots, \left| \frac{\varepsilon^{(0)}(5)}{x^{(0)}(1)} \right| \right) = (0, 0.0326, 0.0359, 0.0279, 0.0275)$$

平均相对误差 $\bar{\Delta} = \frac{1}{n} \sum_{k=1}^n \Delta_k = 0.0248 < 0.05$, 模拟误差 $\Delta_6 = 0.0275 < 0.05$, 精

度为二级。故可用下式进行备件需求量预测:

$$\begin{cases} \hat{x}^{(1)}(k+1) = 4362.8202e^{0.0178k} - 4279.8202 \\ \hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \end{cases}$$

(3) 进行预测。

计算求得预测值 $\hat{x}^{(6)} = 84.1495$, 考虑到实际情况, 在该部队导弹数量及其工作任务量等不变的情况下, 2008 年的需求量取 85 个为宜。图 6-2 为预测值与实际值的对比, 可以清楚地看出预测值误差较小且与实际值趋势大体一致。

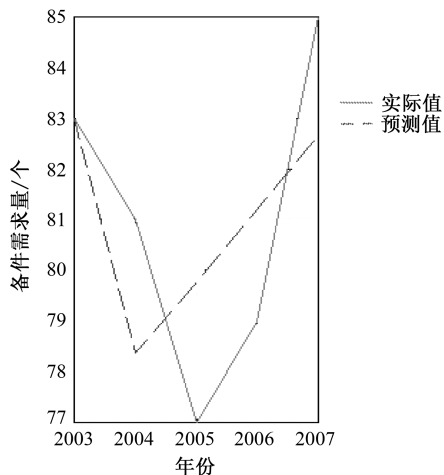


图 6-2 预测值与实际值的对比

6.3 基于含有空缺值序列 GM (1,1) 模型的备件需求预测方法

由于部队战备任务的变更、装备或部件升级换代等原因, 备件消耗量统计数据会因管理因素、人为因素等造成意外的缺失, 使得消耗数据出现间断, 即出现含有空缺值的原始消耗数据序列, 这就给本来数据量不多的备件预测带来更大困难。为充分利用已有的信息, 利用指数平滑模型平滑数据的功能, 提出采用一次指数平滑法生成空缺值, 得到比较光滑的修正序列后建立 GM(1,1)模型, 并进行需求量预测。

6.3.1 模型建立

设空缺序列 $X^{(0)} = (x^{(0)}(1), \dots, x^{(0)}(k-1), \phi(k), \dots, x^{(0)}(n))$, 其中 $\phi(k)$ 为空缺值, 设 $X^{(0)} = (x^{(0)}(1), \dots, x^{(0)}(k-1), x^*(k), \dots, x^{(0)}(n))$ 为修正序列, $x^*(k)$ 为生成值 (修正值)。对于该生成值, 有三种方法。

(1) 非邻均值生成:

$$x^*(k) = 0.5x(k-1) + 0.5x(k+1)$$

生成后, 当且仅当

$$\frac{x(k+1)}{x(k-1)} < \frac{1+\rho(k-1)}{1-\rho(k+1)}$$

时有 $\rho(k-1) < \rho_*(k) < \rho(k+1)$ 。

(2) 灰色生成。根据空缺值的具体位置, 可分别对数列 $X^{(0)} = (x^{(0)}(1), \dots, x^{(0)}(k-1))$ 或 $X^{(0)} = (x^{(0)}(n), x^{(0)}(n-1), \dots, x^{(0)}(k-1))$ 建立 GM (1,1) 模型, 从而求得生成值 $x^*(k)$ 。

(3) 一次指数平滑法生成。设时间序列为 $X^{(0)} = (x^{(0)}(1), \dots, x^{(0)}(k-1), \phi(k))$, 则一次指数平滑公式为

$$S_k^{(1)} = \alpha x_k + (1-\alpha)S_{k-1}^{(1)} \quad (6-9)$$

式中: $S_k^{(1)}$ 为一次指数平滑值; α 为加权系数, 且 $0 < \alpha < 1$ 。

由其预测模型 $\hat{x}_{k+1} = \alpha x_k + (1-\alpha)\hat{x}_k$, 得 $x^*(k) = \phi(k) = \hat{x}_k$ 。

进一步可知, 当 $0 < \alpha < 1$ 且 k 趋向于无穷大时, 指数平滑公式可略写为 $S_k^{(1)} = \alpha \sum_{j=0}^{k-1} (1-\alpha)^j x_{k-j}$ 。显然, $S_k^{(1)}$ 实际上为 $x_k, x_{k-1}, \dots, x_{k-j}, \dots$ 的加权平均。加权系数分别为 $\alpha, \alpha(1-\alpha), \alpha(1-\alpha)^2, \dots$, 有明显指数规律。鉴于上述特点, 且其模型具有平滑数据的功能, 故选用该方法生成空缺值。用此方法时应注重加权系数的选择和初始值的确定。

α 选取的一般原则: 若时间序列波动不大, 则 α 应取小一些, 以减少修正幅度, 使预测模型包含较长时间序列的信息; 若时间序列具有迅速且明显的变化倾向, 则 α 应取大一些, 使预测模型灵敏度高一些, 以便迅速跟上数据的变化。为增强预测精度, 可以多取几个 α 值进行试算, 取预测误差小的 α 值。

$S_0^{(1)}$ 选取的一般原则: 当时间序列数据较多时 (20 以上), 可选用第一期数据为初始值; 当时间序列数据较少 (20 以下), 初始值对以后的预测值影响很大, 可取最初几期实际值的平均值作为初始值。

6.3.2 示例分析

某型导弹侧鳍螺钉 2001—2007 年的消耗量统计数据见表 6-3, 其中 2005 年的消耗量出现空缺, 在预计 2008 年该部队导弹数量及其工作任务等不变的情况下, 预测该型侧鳍螺钉 2008 年的需求量。

表 6-3 某备件 2001—2007 年的消耗数量

年份	2001	2002	2003	2004	2005	2006	2007
备件数量	160	151	170	188	××	150	164

解: 利用一次指数平滑法生成空缺值, 鉴于数据波动较小且数据较少 (少于 20 个), 取 $\alpha = 0.2$, $S_0^{(1)} = \frac{x_1 + x_2}{2} = 155.5$, 得预测值 $\hat{x}_5 = 164.2048$, 根据实际情况

取整数得修正值 $x^*(5) = 165$ 。即对修正序列 $X_1^{(0)} = (160, 151, 170, 188, 165, 150, 164)$ 建立 GM(1,1) 模型并预测, 得到 2008 年该型号侧鳍螺钉需求量为 163(162.9384) 个。

同理, 非邻均值生成与灰色生成的空缺值分别为 169 和 210, 对修正序列 $X_2^{(0)}$ 和 $X_3^{(0)}$ 分别建立 GM(1,1) 模型并进行预测。为体现一次指数平滑法生成空缺值的正确性和优越性, 将三种方法生成的空缺值、对修正序列建立的 GM(1,1) 模型的平均相对误差、精度等级及预测结果做比较, 见表 6-4。从表 6-4 中可明显看出, 对修正序列 $X_1^{(0)}$ 建立的 GM(1,1) 模型具有较为满意的精度和可信度。而

对修正序列 $X_2^{(0)}$ 和 $X_3^{(0)}$ 建立的 GM(1,1)模型的精度等级为三级,严格意义上需要经过模型修正才能用其进行预测。为此,会使整个预测过程变得比较复杂。所以,选用一次指数平滑法生成空缺值能得到比较光滑的修正序列,有利于建立精度等级良好、可以直接用于预测的 GM(1,1)模型。

表 6-4 预测效果比较

空缺值生成方法	空缺值	平均相对误差	精度等级	2008 年预测值
一次指数平滑生成	165	0.0498	二级	163
非邻均值生成	169	0.0526	三级	161
灰色生成	210	0.0878	三级	175

6.4 基于加入影响因子改进灰色模型的备件需求预测方法

6.4.1 改进灰色预测模型的建立

运用灰色 GM(1,1)模型进行备件消耗量预测时,如果备件消耗受任务、环境、人员操作等因素影响较大,呈现较大波动趋势,此时通过经典的灰色预测模型很难达到理想的效果。

1. 在预测模型中添加影响因子

考虑级比和影响因子,得出模型改进公式

$$\begin{cases} x^{(1)}(k) = x^{(1)}(k) (k = 1, 2, \dots, n-1) \\ x^{(1)}(n) = (2 + a - \sigma(n))x^{(1)}(n) \end{cases} \quad (6-10)$$

式中: $x^{(1)}(k) = x^{(1)}(k) (k = 1, 2, \dots, n-1)$ 。式 (6-10) 代表在进行修正时相对于 n 前的 $n-1$ 个数不变,只是在第 n 个数据的值上加以修正; $\sigma(n)$ 为 x 在 n 时刻的级比; a 为预测时刻数据的估计百分比,采用处理后的序列建立 GM(1,1)模型,采用这种方法建立的模型即加入影响因子的 GM(1,1)模型。

2. 影响因子的确定

以往给出的影响因子的判别式很笼统,或是指出依照上级命令,给出 a 值,

或是说受到某一因素的影响， a 值应该是多少之类，这些说辞没有任何的依据，给出的结果很值得人们进行推敲，同时让他人在进行相关预测时，感到无从下手。本书在试验分析的基础上，给出了自己的理解，分析过程如下：

(1) 分析备件消耗的影响因素。

(2) 对影响因素进行量化，得到样本数据。描述某因素某年量化后的数据 $a_t(n)$ ，其中 t 为年度， n 为该影响因素的序号。

(3) 对各影响因素进行无量纲化处理。在获得样本数据后，由于其中各个指标因素互不相同，原始样本中各量的单位与量级差别也不同，需要对样本数据进行无量纲化处理，将数据处理为区间 $[0,1]$ 之间的数据。无量纲化处理方法很多，这里采用公式

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (6-11)$$

设 $a_t(n)$ 在处理后的值为 $\hat{a}_t(n)$ 。

(4) 求出影响因子。影响因子为

$$a = \pm \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\hat{a}_t(i) - \bar{\hat{a}}_t(i))^2} \quad (6-12)$$

式中： $\bar{\hat{a}}_t(n) = \frac{1}{n} \sum_{i=1}^n \hat{a}_t(i)$ 。

这里影响因子的值可正可负。经反复提取数据进行运算验证，可得出判断准则如下：

在某一年度，比较 $\hat{a}_t(n)$ 与 $\bar{\hat{a}}_t(n)$ ，设有 k 组情况满足 $\bar{\hat{a}}_t(n) \geq \hat{a}_t(n)$ ，当 $k < \frac{n}{2}$

时， a 取正值，否则取负值。

当在预测过程中没有重大因素的影响时，就不必加入影响因子，直接预测即可满足需要。

6.4.2 示例分析

螺栓是某型导弹装备的常用零部件，其 1998—2007 年的消耗量统计数据如表 6-5 所列，试确定该部件备件的消耗规律。

1. 利用 GM(1,1)模型，计算预测值并与实际值进行比较

原始序列为 $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(10)\} = \{40, 34, 20, \dots, 36\}$

表 6-5 螺栓 1998—2007 年的消耗量

序号	1	2	3	4	5	6	7	8	9	10
年份	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
消耗量/件	40	34	20	25	35	42	38	33	24	36

由图 6-3 可以看出，原始序列呈不规则的波动曲线，其含义随机变量很难找出变化规律，也难以用解析函数来拟合。

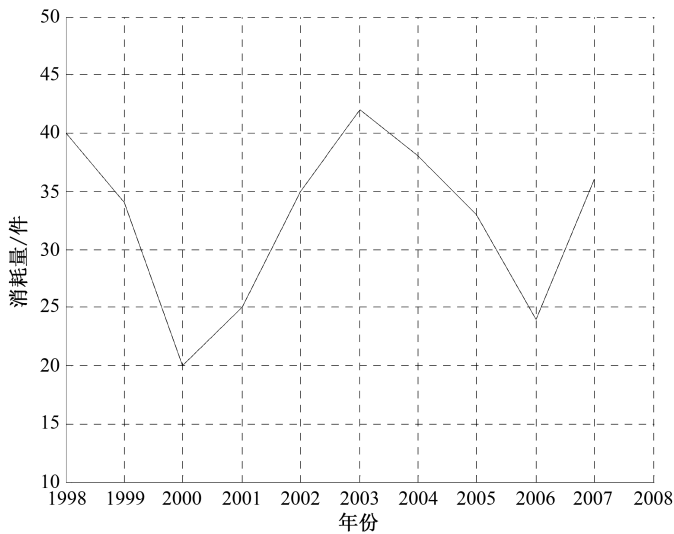


图 6-3 消耗量曲线

对表原始数据做累加生成处理得到新的数据列：

$$X_k^{(1)} = \sum_{m=1}^k X_m^{(0)}$$

累加后的数据列为

$$X_{10}^{(1)} = \{40 \quad 74 \quad 94 \quad 119 \quad 154 \quad 196 \quad 234 \quad 267 \quad 291 \quad 327\}$$

经过一次累加生成以后形成的序列，其曲线如图 6-4 所示。

可以看出，新序列 $X^{(1)}$ 具有单调增加的趋势，可以用适当的函数来描述。经检验该序列具有光滑性，符合灰色模型应用条件，因此可以运用灰色模型系统进行建模。代入 GM(1,1)模型，利用模型求解参数 a 与 b 的值，计算过程如下：

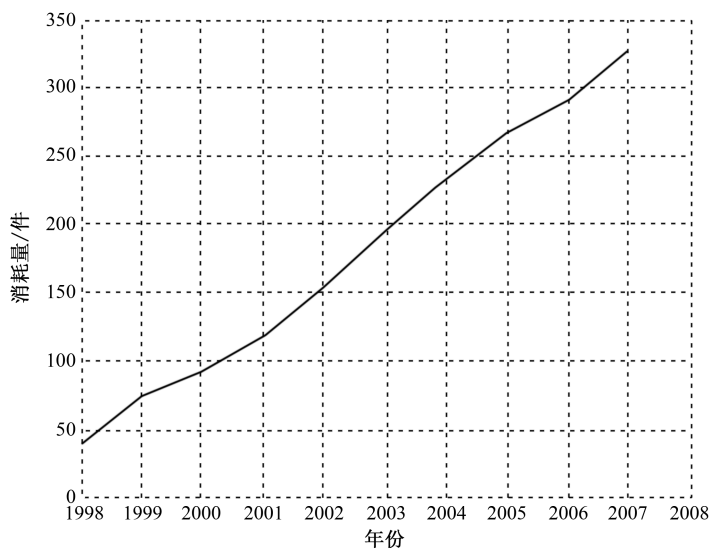


图 6-4 累加后的消耗量曲线

$$B = \begin{bmatrix} -57.0 & 1 \\ -84.0 & 1 \\ -106.5 & 1 \\ -136.5 & 1 \\ -175.0 & 1 \\ -215.0 & 1 \\ -250.5 & 1 \\ -279.0 & 1 \\ -309.0 & 1 \end{bmatrix}$$

$$B^T = \begin{bmatrix} -57.0 & -84.0 & -106.5 & -136.5 & -175.0 & -215.0 & -250.5 & -279.0 & -309.0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$Y_n = \{34 \ 20 \ 25 \ 35 \ 42 \ 38 \ 33 \ 24 \ 36\}^T$$

将数据代入下式:

$$C = (B^T B)^{-1} B^T Y_n = \begin{pmatrix} a \\ b \end{pmatrix}$$

其结果如下:

$$C = (B^T B)^{-1} B^T Y_n = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -0.0193443 \\ 28.423351 \end{pmatrix}$$

$$a = -0.0193443, \quad b = 28.423351, \quad b/a = -1469.3$$

$$X_1^{(1)} - \frac{b}{a} = 1509.3$$

将结果中求得的 a 与 b 值代入, 可得

$$\hat{X}_{k+1}^{(1)} = 1509.3e^{0.0193443k} - 1469.3$$

此时求出的值为累加以后的值, 要得到预测的年度备件消耗量必须对此预测数列进行还原, 将计算出的数据进行还原, 即可得年度预测值。实际值与预测值的对比如表 6-6 所列, 实际值与预测值的曲线如图 6-5 所示。

表 6-6 实际值与预测值对比

年份	实际值	预测值	残差	误差/%
1998	40	—	—	—
1999	34	29.5	4.5	13.2
2000	20	30.1	-10.1	-50.5
2001	25	30.6	-5.6	-22.4
2002	35	31.2	3.8	10.9
2003	42	31.8	10.2	24.3
2004	38	32.5	5.5	14.5
2005	33	33.1	-0.1	-0.3
2006	24	33.8	-9.8	-40.8
2007	36	34.4	1.6	4.4

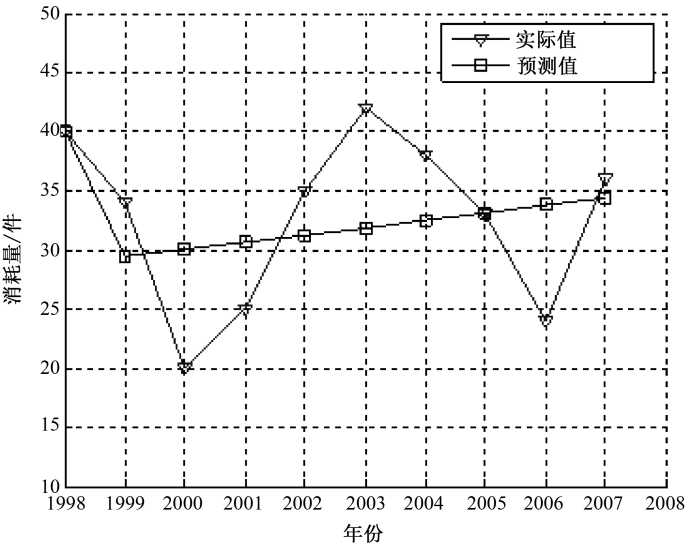


图 6-5 消耗量的实际值与预测值曲线图

通过原始数据与预测数据对比,以及备件实际消耗值与预测消耗值对比,可以看出直接套用 GM(1,1)模型而不加任何修正条件,得到的预测数据与实际数据有很大的差别。2000 年,预测数量与实际发出备件数量相比残差绝对值达 10.1%之多,可见预测数据与实际数据有很大的差别。

2. 利用影响因子,改进 GM(1,1)模型

1) 分析影响备件消耗的因素并量化

经分析影响该备件消耗的因素主要有 6 个,列出如下:

A——备件的固有可靠性,把导弹装备定检过程中发现的失效零部件数作为自然消耗的备件数并对其量化。

B——环境因素,根据该部队气象部门的统计资料,将其量化为根据环境的严酷度所划分的 7 个等级,并相应地量化为 1~7 范围的整数值,即数值越大表示严酷程度越高,环境越恶劣。

C——人为因素,以导弹技术准备过程中由于人为因素导致的零部件失效数为该影响因素的量化值。

D——年度重大演习任务次数,直接将次数作为量化值。

E——执行任务时间,根据平时的训练计划可以统计出该型号导弹年度战备值班的总时间,单位为小时,将其作为量化值。

F——年度使用、维修与定检中该部件被拆卸的总次数,将总次数作为量化值。

6 个影响因素的相关数据及历史消耗数据如表 6-7 所列。

表 6-7 影响因素量化值

年份	影响因素量化值						实际消耗量
	A	B	C	D	E	F	
1998	2	5	3	2	750	34	40
1999	0	4	4	1	450	24	34
2000	1	3	3	0	460	20	20
2001	0	2	1	1	360	28	25
2002	0	3	4	0	560	36	35
2003	1	3	2	2	880	30	42
2004	2	5	2	3	780	38	38
2005	2	1	0	1	700	30	33
2006	0	1	2	2	760	42	24

2007	3	2	1	1	570	45	36
------	---	---	---	---	-----	----	----

2) 对各影响因素进行无量纲化处理
计算后, 处理结果如表 6-8 所列。

表 6-8 影响因素处理

年份	影响因素量化值						实际 消耗量
	A	B	C	D	E	F	
1998	0.6667	1.0000	0.7500	0.6667	0.7500	0.5600	40
1999	0.0000	0.7500	1.0000	0.3333	0.1731	0.1600	34
2000	0.3333	0.5000	0.5000	0.0000	0.1923	0.0000	20
2001	0.0000	0.2500	0.2500	0.3333	0.0000	0.3200	25
2002	0.0000	0.5000	1.0000	0.0000	0.3846	0.6400	35
2003	0.3333	0.5000	0.5000	0.6667	1.0000	0.4000	42
2004	0.6667	1.0000	0.5000	1.0000	0.8077	0.7200	38
2005	0.6667	0.0000	0.0000	0.3333	0.6538	0.4000	33
2006	0.0000	0.0000	0.5000	0.6667	0.7692	0.8800	24
2007	1.0000	0.2500	0.2500	0.3333	0.4038	1.0000	36

3) 分析备件消耗的影响因子
1999 年, $\bar{\hat{a}}_t(n)$ 的计算如下:

$$\begin{aligned}\bar{\hat{a}}_t(n) &= \frac{1}{n} \sum_{i=1}^n \hat{a}_t(i) = \frac{1}{6} (0.0000 + 0.7500 + 1.0000 + 0.3333 + 0.1731 + 0.1600) \\ &= 0.402733\end{aligned}$$

比较 $\hat{a}_t(n)$ 与 $\bar{\hat{a}}_t(n)$ 可知, 满足 $\bar{\hat{a}}_t(n) \geq \hat{a}_t(n)$ 的 k 值为 2, 有 $k = 2 < 3 = \frac{n}{2}$, a 应取正值。

2000 年, $\bar{\hat{a}}_t(n)$ 的计算如下:

$$\begin{aligned}\bar{\hat{a}}_t(n) &= \frac{1}{n} \sum_{i=1}^n \hat{a}_t(i) = \frac{1}{6} (0.3333 + 0.5000 + 0.5000 + 0.0000 + 0.1923 + 0.0000) \\ &= 0.254267\end{aligned}$$

比较 $\hat{a}_t(n)$ 与 $\bar{\hat{a}}_t(n)$ 可知, 满足 $\bar{\hat{a}}_t(n) \geq \hat{a}_t(n)$ 的 k 值为 4, 有 $k = 4 > 3 = \frac{n}{2}$, a 应取负值。

利用

$$a = \pm \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\hat{a}_t(i) - \bar{\hat{a}}_t(i))^2}, \quad \bar{\hat{a}}_t(n) = \frac{1}{n} \sum_{i=1}^n \hat{a}_t(i)$$

取得的影响因子见表 6-9。

表 6-9 影响因子

年份	1998	1999	2000	2001	2002
影响因子 a	-0.148717	0.388841	-0.228105	-0.152843	0.386070
年份	2003	2004	2005	2006	2007
影响因子 a	0.240379	0.143703	-0.296653	-0.384450	0.361296

4) 计算加入影响因子后的预测值

$\sigma(n) = \frac{x_n^{(1)}}{x_{n-1}^{(1)}}$ 为 x 在 n 时刻的级比。为获得精确的预测结果，并与影响因子的

数值相协调，这里级比数值的小数点后取六位有效数字。

$$\sigma(2) = \frac{x_2^{(1)}}{x_1^{(1)}} = \frac{29.5}{40} = 0.737500$$

同理，可求出全部的级比，见表 6-10。

表 6-10 级比值

级比	$\sigma(2)$	$\sigma(3)$	$\sigma(4)$	$\sigma(5)$	$\sigma(6)$
数值	-0.148717	0.388841	-0.228105	-0.152843	0.386070
级比	$\sigma(7)$	$\sigma(8)$	$\sigma(9)$	$\sigma(10)$	—
数值	0.240379	0.143703	-0.296653	-0.384450	—

由表 6-11 可看出各种因素对备件消耗量的制约，假如不考虑这几个因素的制约，预测数值将比实际值有很大的差别。由此可见，考虑其他影响因子的重要性，尤其作为备件的使用，它与上级作战命令及保障任务的多少是分不开的，同时也受到天气的影响。备件保障中各种临时因素的制约为工作增加了难度，但是影响因子的引入，为我们及时准确地进行备件预测工作起到了重要的作用。根据不同时刻各种影响因子的作用，可以看出预测的准确度得到了有效保障。由图 6-6 可以看出加入影响因子的作用，预测值的精确度有了很大的提高，这种模型为我们进行备件保障提供了很好的依据。

表 6-11 实际值与加入影响因子后的预测值对照

年份	实际值	加入影响因子后的预测值	残差	误差/%
1998	40	—	—	—
1999	34	48.7	-4.7	-13.8
2000	20	22.6	-2.6	-13.0
2001	25	25.4	-0.4	-1.6
2002	35	42.6	-7.6	-21.7
2003	42	38.8	3.2	7.6
2004	38	36.5	1.5	3.9
2005	33	22.7	10.3	31.2
2006	24	20.1	3.9	16.3
2007	36	46.2	-10.2	45.0

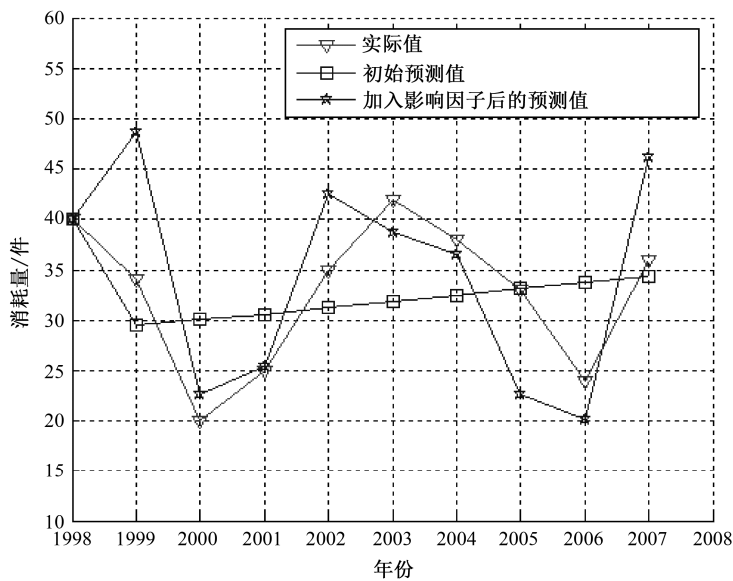


图 6-6 消耗量的总对比

3. 模型的检验

对上述模型进行后验差检验，其计算如下：

$X^{(0)}$ 的均值为

$$\bar{X} = \frac{1}{10} \sum_{i=1}^{10} X^{(0)}(10) = 32.7$$

$X^{(0)}$ 的方差为

$$S_1^2 = \frac{1}{n} \sum_{k=1}^n (x^{(0)}(k) - \bar{x})^2 = 48.21$$

残差的均值为

$$\bar{\varepsilon} = \frac{1}{n} \sum_{k=1}^n \varepsilon(k) = -0.73$$

残差的方差为

$$S_2^2 = \frac{1}{n} \sum_{k=1}^n (\varepsilon(k) - \bar{\varepsilon})^2 = 9.96$$

均方差比值为

$$C = \frac{S_2}{S_1} = 0.455 \quad (\text{模型精度为 1 级})$$

小误差概率 $p = 0.7$ ，模型精度为 2 级。

6.5 基于灰色-马尔可夫模型的备件需求预测方法

在实际应用中，由于原始数据个数有限且具有起伏性和无序性，灰色预测模型通常难以将预测带限制在一个较小的范围内，导致预测精度不高。另外，GM(1,1)模型还必须要求原始序列为非负准光滑序列（若不能满足这个条件，则需对原始数列进行预处理），这样所累加生成的数据序列才具有指数性质；否则，将产生较大误差，而由一个时间序列所累加生成的数据序列未必满足这个条件。这些问题可以通过对灰色预测模型的结果进行马尔可夫链改进来解决。事实上，作为灰色组合模型形式之一的灰色马尔可夫预测模型具有良好的适应能力，有效地提高了具有随机波动性数据的预测精度。

6.5.1 经典灰色-马尔可夫预测模型

把灰色预测与马尔可夫链预测两者结合起来，所形成的灰色-马尔可夫预测模型不但能够用于揭示数据序列的发展趋势，而且能够用于预测状态变化规律，使得预测具有较好的科学性和实用性。

1. 经典灰色-马尔可夫预测模型的计算过程

设数据序列 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ ，利用灰色-马尔可夫预测模型针对该数据序列进行数据预测的计算过程如下：

(1) 建立灰色 GM(1,1)模型，求出其预测曲线 $\gamma'(k) = \hat{x}^{(0)}(k+1)$ 。

(2) 状态划分：以 $\gamma'(k)$ 曲线为基准，分割成与 $\gamma'(k)$ 曲线平行的数个条形区域，这样的每个条形区域就构成马尔可夫链中的某一个状态，任一状态区间表示为

$$\otimes_i = [s_{1i}, s_{2i}] (i=1, 2, \dots, L)$$

式中

$$s_{1i} = \gamma'(k) + A_i, \quad s_{2i} = \gamma'(k) + B_i$$

其中： A_i 、 B_i 分别为上限的偏移，可酌情设置大小。

(3) 计算转移概率矩阵：

$$P_{ij}(m) = M_{ij}(m) / M_i (i, j=1, 2, 3, \dots, L)$$

式中： $P_{ij}(m)$ 为由状态 \otimes_i 经过 m 步转移到 \otimes_j 的概率； L 为所划分的状态总数； M_i 为原始数据按一定概率落入状态 \otimes_i 的样本数； $M_{ij}(m)$ 为由状态 \otimes_i 经 m 步转移到 \otimes_j 的样本数。

由此，可得出马尔可夫状态转移概率矩阵为

$$P(m) = \begin{bmatrix} p_{11}(m) & \cdots & p_{1L}(m) \\ \vdots & & \vdots \\ p_{L1}(m) & \cdots & p_{LL}(m) \end{bmatrix}$$

下面考查 $P(1)$ 中的第 k 行，若 $\max_j p_{kj} = p_{kd}$ ，则下一时刻转移到 \otimes_d 状态。

如果使用一步转移概率矩阵 $P(1)$ 难以确定状态的未来转向，则需要考查多步转移概率矩阵 $P(m)$ 。

(4) 确定预测值：当 \otimes_d 确定后，也就确定了预测值的变动区间 $[s_{1d}, s_{2d}]$ ，则有

$$\hat{\gamma}(k) = \gamma'(k) + \frac{A_i + B_i}{2}$$

(5) 模型精度检验：

$$q^{(0)}(k) = x^{(0)}(k) - \hat{x}^{(0)}(k+1), \quad \varepsilon(k) = q^{(0)}(k) / x^{(0)}(k)$$

式中： $q^{(0)}(k)$ 为残差； $\varepsilon(k)$ 为相对误差。

2. 灰色-马尔可夫链的状态划分问题

灰色-马尔可夫链的状态划分通常有相对值法、平行曲线法、残差的标准化离差法、相对误差法等多种方法。下面以常用的平行曲线法和残差的标准化离差法为例说明灰色-马尔可夫链的状态划分的过程。

平行曲线法将预测值 $\hat{X}^{(0)}$ 序列绘制成一光滑的曲线 z ，形成预测曲线，将预测曲线附近的区域划分成与预测值曲线平行的若干区间 \otimes_i ，每一区间构成一个状态： $\otimes_i \in [\hat{x}^{(0)}(1 + \Delta_{1i}), \hat{x}^{(0)}(1 + \Delta_{2i})]$ ， Δ_{1i} 和 Δ_{2i} 可根据具体情况自行确定，大小无严格要求，但当数据量较小时以划分较少的状态为宜，以保证预测的精确。平行曲线法状态划分如图 6-7 所示。

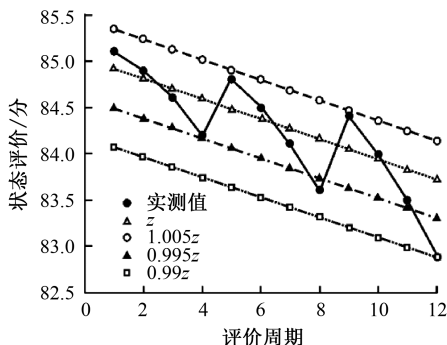


图 6-7 平行曲线法状态划分

构成与预测曲线平行的三个状态区间分别是 $\otimes_1[z, (1+0.005)z]$ 、 $\otimes_2[(1-0.005)z, z]$ 和 $\otimes_3[(1-0.01)z, (1-0.005)z]$ ，样本的 12 个预测值完全包含于三个状态区间中，即对应的灰色-马尔可夫链可分为三个状态。

划分状态是马尔可夫预测中的关键性步骤。划分状态区间时，一是要将预测值完全覆盖到，二是状态区间不能两两重叠，三是状态区间的大小和位置要合适。一般来说，原始数据较少时，划分区间宜少，以便增多各状态间的转移次数，从而更加客观地反映系统各状态间的转移规律；原始数据较多时，区间不妨划分多一些，以便从历史数据中挖掘更多的信息，以提高预测精度。此外，许多文献划分的状态区间一般是在预测曲线附近、与预测值曲线平行的若干宽度相等的区间。

可见，状态区间的划分具有较大的随意性，这导致预测的精确度无法总是得到保证。

6.5.2 改进的灰色-马尔可夫预测模型

在上述基于灰色-马尔可夫预测模型的一般预测过程中, 状态转移概率 P_{ij} 的计算是该预测方法的核心。传统的状态转移概率计算方法建立在一个实际数据对应一个确定的状态的基础上。有时候会出现这样的情况: 某一个实际数据恰好处于状态区间的分界线上, 数据属于“亦此亦彼”状态。对于这种情况, 传统的马尔可夫链状态转移概率将无法计算。为了避免这种情况, 必须在确定状态划分时就有意使状态分界线“绕开”任何数据, 但是, 这实际上增加了状态划分的“主观随意性”。

另外, 从常理判断, 属于某状态区间但处于状态区间分界线附近的数据, 与处于该状态区间中间的数据相比, 它们“属于”该状态的“程度”应该是有区别的。只有更加精确地描述数据的“状态隶属程度”, 才能更精确地描述数据在状态之间的变化规律。显然, 建立在一个实际数据对应一个确定的状态基础上的传统状态划分方法没有这种描述能力。

基于上述分析, 提出一种新的状态转移概率计算方法, 而这种方法的基础是状态隶属度的概念及计算。

1. 状态隶属度的概念

在状态划分确定之后, 当数据处于两个相邻状态区间的中线之间时, 每个实际数据不完全属于某一个确定的状态。该数据分别以一定的隶属度属于这两个状态, 即每个数据针对不同的状态具有一定的状态隶属度。

具体地, 当数据 x 处于状态区间 \otimes_i 的中线上时, 该数据对该状态的隶属度为 1; 当数据 x 处于状态区间 \otimes_i 的中线和状态区间 \otimes_{i+1} 的中线之间时, x 分别以隶属度 $A_i(x)$ 和 $A_{i+1}(x)$ 属于两个状态 \otimes_i 和 \otimes_{i+1} 。其中, $\otimes_i = [s_{1i}, s_{2i}]$, s_{1i} 、 s_{2i} 分别为区间 \otimes_i 的下界值和上界值。

状态隶属度的大小可根据数据 x 与两个相邻状态区间中线的距离来衡量:

$$A_i(x) = 1 - \frac{|x - \bar{s}_i|}{|\bar{s}_i - \bar{s}_{i+1}|} \quad (6-13)$$

$$A_{i+1}(x) = \frac{|x - \bar{s}_i|}{|\bar{s}_i - \bar{s}_{i+1}|} \quad (6-14)$$

式中: \bar{s}_i 、 \bar{s}_{i+1} 分别为状态 \otimes_i 和 \otimes_{i+1} 区间的中间值, 且 $\bar{s}_i \neq \bar{s}_{i+1}$ 。

显然, 某一个数据 x 的状态隶属度 $A_i(x)$ 和 $A_{i+1}(x)$ 满足

$$A_i(x) + A_{i+1}(x) = 1 \quad (6-15)$$

2. 状态隶属度的计算

设已划分 L 个状态区间为 $\otimes_1, \otimes_2, \dots, \otimes_i, \otimes_{i+1}, \dots, \otimes_L$, 与状态区间 \otimes_i 相邻的状态区间为 \otimes_{i-1} 和 \otimes_{i+1} 。考查状态区间 \otimes_i 内的数据 x , 其隶属度的计算分几种情况: 数据 x 隶属于 \otimes_{i-1} 和 \otimes_i ; 数据 x 隶属于 \otimes_i 和 \otimes_{i+1} ; 数据 x 只隶属于 \otimes_1 或 \otimes_L 。不妨假设 $\bar{s}_1 < \bar{s}_2 < \dots < \bar{s}_L$, 即状态区间中间值与状态序号成正比。

情况一: x 隶属于 \otimes_{i-1} 和 \otimes_i 。

对于状态区间 \otimes_i 内的数据 x , 当 $x < \bar{s}_i$ 时, 数据 x 隶属于 \otimes_{i-1} 和 \otimes_i ($2 \leq i < L$)。

x 对状态 \otimes_{i-1} 的隶属度为

$$A_{i-1}(x) = \frac{|x - \bar{s}_i|}{|\bar{s}_i - \bar{s}_{i-1}|} \quad (6-16)$$

x 对状态 \otimes_i 的隶属度为

$$A_i(x) = 1 - \frac{|x - \bar{s}_i|}{|\bar{s}_i - \bar{s}_{i-1}|} \quad (6-17)$$

情况二: x 隶属于 \otimes_i 和 \otimes_{i+1} 。

对于状态区间 \otimes_i 内的数据 x , 当 $x > \bar{s}_i$ 时, 数据 x 隶属于 \otimes_i 和 \otimes_{i+1} ($2 \leq i < L$)。

x 对状态 \otimes_i 的隶属度为

$$A_i(x) = 1 - \frac{|x - \bar{s}_i|}{|\bar{s}_i - \bar{s}_{i+1}|} \quad (6-18)$$

x 对状态 \otimes_{i+1} 的隶属度为

$$A_{i+1}(x) = \frac{|x - \bar{s}_i|}{|\bar{s}_i - \bar{s}_{i+1}|} \quad (6-19)$$

情况三: x 只隶属于 \otimes_1 , 特别地, 对于状态区间 \otimes_1 内的数据 x , 当 $x < \bar{s}_1$ 时, 数据 x 完全隶属于 \otimes_1 , 即 $A_1(x) = 1$ 。

情况四: x 只隶属于 \otimes_L , 特别地, 对于状态区间 \otimes_L 内的数据 x , 当 $x > \bar{s}_L$ 时, 数据 x 完全隶属于 \otimes_L , 即 $A_L(x) = 1$ 。

3. 状态转移概率的确定

下面以残差的标准化离差法为例, 说明传统马尔可夫链状态转移概率和改进的马尔可夫链状态转移概率的确定方法。相对值法、平行曲线法和相对误差法可基于该方法类似扩展可得。

传统马尔可夫链状态转移概率的确定方法如下:

设对残差相对值序列进行状态划分, 分为 L 个状态, 系统所处的任一状态记为 $\otimes_i (1 \leq i \leq L)$, 状态区间 $\otimes_i = [s_{1i}, s_{2i}]$, 其中, s_{1i} 、 s_{2i} 分别为状态 \otimes_i 的下界值和上界值。设系统由状态 \otimes_i 转移到状态 \otimes_j 的数据样本数为 M_{ij} , 系统处于状态 \otimes_i 的数据样本数为 M_i , 则转移概率 $p_{ij} = \frac{M_{ij}}{M_i}$, 显然有

$$\sum_{j=1}^L p_{ij} = \frac{\sum_{j=1}^L M_{ij}}{M_i} = 1 \quad (p_{ij} \geq 0) \quad (6-20)$$

在传统方法的基础上, 改进的马尔可夫链状态转移概率确定方法如下:

采用状态隶属度对状态转移的数据样本数进行修正, 进而改进状态转移概率的计算方法。下面根据状态隶属度不同的计算方法给出相应的状态转移数据样本数计算方法。

考查数据 x_0 在 t 时刻处于状态 \otimes_i , $t+1$ 时刻转移到状态 \otimes_j 的情况。按照传统马尔可夫链状态转移数据样本的计算方法, x_0 在 t 时刻处于状态 \otimes_i , $t+1$ 时刻转移到状态 \otimes_j 的转移次数为 1。按照改进的马尔可夫链状态转移数据样本计算方法, 设数据 x_0 在 t 时刻处于状态 \otimes_i , $t+1$ 时刻转移到状态 \otimes_j 时, 转移到 x , $x \in \otimes_j$, x 并不一定完全隶属于状态 \otimes_j , 还可能部分隶属于 \otimes_{j-1} 和 \otimes_{j+1} 。

下面用状态隶属度来描述数据 x_0 的一次转移次数, 分以下几种情况 (设数据转移到 x , $x \in \otimes_j$):

(1) $\bar{s}_{j-1} < x < \bar{s}_j$ 。数据 x 在 t 时刻处于状态 \otimes_i , $t+1$ 时刻转移到状态 \otimes_j 的状态转移数为

$$M_{ij}(t) = A_j(x) = 1 - \frac{|x - \bar{s}_j|}{|\bar{s}_j - \bar{s}_{j-1}|} \quad (6-21)$$

转移到状态 \otimes_{j-1} 的状态转移数为

$$M_{i,j-1}(t) = A_{j-1}(x) = \frac{|x - \bar{s}_j|}{|\bar{s}_j - \bar{s}_{j-1}|} \quad (6-22)$$

因此, 一次状态转移总数为

$$M_{ij}(t) + M_{i,j-1}(t) = A_j(x) + A_{j-1}(x) = 1 \quad (6-23)$$

(2) $\bar{s}_j < x < \bar{s}_{j+1}$ 。类似地, 数据 x 在 t 时刻处于状态 \otimes_i , $t+1$ 时刻转移到状

态 \otimes_j 的状态转移数确定为

$$M_{ij}(t) = A_j(x) = 1 - \frac{|x - \bar{s}_j|}{|\bar{s}_j - \bar{s}_{j-1}|} \quad (6-24)$$

转移到状态 \otimes_{j+1} 的状态转移数为

$$M_{i,j+1}(t) = A_{j+1}(x) = \frac{|x - \bar{s}_j|}{|\bar{s}_j - \bar{s}_{j+1}|} \quad (6-25)$$

一次状态转移总数为

$$M_{ij}(t) + M_{i,j+1}(t) = A_j(x) + A_{j+1}(x) = 1 \quad (6-26)$$

(3) x 只隶属于 \otimes_1 。如果数据 x 在 t 时刻处于状态 \otimes_i ， $t+1$ 时刻转移到状态 \otimes_1 ，且 $x < \bar{s}_1$ 时，数据 x 完全隶属于 \otimes_1 ，则一次状态转移数为

$$M_{i1}(t) = A_1(x) = 1 \quad (6-27)$$

(4) x 只隶属于 \otimes_L 。如果数据 x 在 t 时刻处于状态 \otimes_i ， $t+1$ 时刻转移到状态 \otimes_L ，且 $x > \bar{s}_L$ 时，数据 x 完全隶属于 \otimes_L ，则一次状态转移数为

$$M_{iL}(t) = A_L(x) = 1 \quad (6-28)$$

综上所述，采用改进的状态转移概率计算方法，由于一次转移次数等值为状态隶属度之和，因此，状态 \otimes_i 转移到状态 \otimes_j 的每个状态转移数不一定为非负整数。

虽然这样处理后，单个状态转移概率值发生变化，但通过以上分析可知， x 在 t 时刻处于状态 \otimes_i ， $t+1$ 时刻转移到状态 \otimes_j 时， $M_{ij}(t) + M_{i,j-1}(t) = 1$ 或

$M_{ij}(t) + M_{i,j+1}(t) = 1$ 成立，一次转移总次数仍恒为1， $\sum_{j=1}^L M_{ij}$ 的值与传统计算方法

法中 $\sum_{j=1}^L M_{ij}$ 的数值相等，即状态转移总数 M_{ij} 与传统计算方法相比不变，仍满足

$$\sum_{j=1}^L p_{ij} = \frac{\sum_{j=1}^L M_{ij}}{M_i} = 1 \text{ 的条件。}$$

4. 基于改进的状态转移概率矩阵的预测步骤

(1) 根据原始数据 $x(k)$ ，求出GM(1,1)模型中的系数 a 、 b ，得到与实际数据对应的拟合值 $\hat{x}(k)$ 。

(2) 由残差相对值公式

$$\varepsilon(k) = \frac{x(k) - \hat{x}(k)}{\hat{x}(k)}$$

求得残差相对值序列 $\varepsilon(k)$ 的范围。

(3) 根据实际情况将残差相对值范围划分为 L 个状态: $\otimes_1, \otimes_2, \dots, \otimes_L$ 。

(4) 由改进的马尔可夫链状态转移概率确定方法, 计算一步状态转移概率矩阵:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1L} \\ p_{21} & p_{22} & \cdots & p_{2L} \\ \vdots & \vdots & & \vdots \\ p_{L1} & p_{L2} & \cdots & p_{LL} \end{pmatrix}$$

式中: $p_{ij} \geq 0$, $\sum_{j=1}^L p_{ij} = 1$ 。

假设一个预测系统有 \otimes_1 、 \otimes_2 、 \otimes_3 、 \otimes_4 四个状态, 实际数据 $x(k)$ 中最后一个数据处于状态 \otimes_4 , 则初始分布为 $I^{(0)} = (0 \ 0 \ 0 \ 1)$, 下一个数据预测值为 $I^{(1)} = I^{(0)} \cdot P$, 据此可得出预测数据所处的状态区间。

6.5.3 示例分析

海军航空兵某飞行团参加了 8 次索马里护航, 某航材备件消耗的历史数据如表 6-12 所列。

表 6-12 某航材备件消耗的历史数据

序号	1	2	3	4	5	6	7	8
备件数量	20	14	18	9	10	11	6	5

下面以 1~7 次的数据为样本, 分别以改进的灰色-马尔可夫模型和其他模型预测第 8 次的备件数量, 并做对比分析。

1. 数据的检验与处理方法

设参考数据为 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, 针对数据序列的级比:

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)} \quad (k = 2, 3, \dots, n)$$

如果所有的级比 $\lambda(k)$ 都落在可容覆盖 $\Phi = (e^{-\frac{2}{n+1}}, e^{\frac{2}{n+2}})$ 内, 则数据序列 $x^{(0)}$

为非负准光滑序列。

否则,为使数据序列 $x^{(0)}$ 落入可容覆盖内,取适当的常数 c ,做平移变换 $y^{(0)}(k) = x^{(0)}(k) + c (k=1, 2, \dots, n)$,使数列 $Y^{(0)} = (y^{(0)}(1), y^{(0)}(2), \dots, y^{(0)}(n))$ 的级比 $\lambda_y(k) = \frac{y^{(0)}(k-1)}{y^{(0)}(k)} \in \Phi (k=2, 3, \dots, n)$ 。

2. 预测误差估计方法

单点相对误差计算:

$$e(k) = \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \times 100\% (k=2, 3, \dots, n)$$

平均误差计算:

$$\bar{E} = \frac{\sum_{k=2}^n |e(k)|}{n-1}$$

3. 求解过程

(1) 已知 $X^{(0)} = \{20, 14, 18, 9, 10, 11, 6, 5\}$, 可得

$$\lambda_x = \{1.4268, 0.7778, 2.0000, 0.9000, 0.9091, 1.8333, 1.2000\}$$

而 $\Theta = (0.8007, 1.2214)$, 所以需要对 $X^{(0)}$ 进行平移变换。取 $c = 32.5$, 可得

$$Y^{(0)} = \{52.5, 46.5, 50.5, 41.5, 42.5, 43.5, 38.5, 37.5\}$$

$$\lambda_y = \{1.1290, 0.9208, 1.2169, 0.9765, 0.9770, 1.1299, 1.0267\}$$

于是, $Y^{(0)}$ 序列可以进行 GM(1,1) 模型运算, 最后再减去 c 可得到 $\hat{X}^{(0)}$ 。

(2) 对 $Y^{(0)}$ 做 1-AGO, 可得

$$Y^{(1)} = \{52.5, 99, 149.5, 191, 233.5, 277.5, 315.5, 353\}$$

(3) 对 $Y^{(1)}$ 用紧邻均值生成序列 Z

$$Z^{(1)} = \{75.75, 124.25, 170.25, 212.25, 255.25, 296.25, 334.25\}$$

(4) 建立预测方程

$$\hat{y}^{(1)}(k) = -1216.14 \times e^{-0.0405412 \times (k-1)} + 1268.64$$

(5) 求出还原值

$$\hat{Y}^{(0)} = \{52.5, 48.3176, 46.398, 44.5545, 42.7844, 41.0845, 39.4522, 37.8848\}$$

所以

$$\hat{X}^{(0)} = \{20.0000, 15.8176, 13.8980, 12.0545, 10.2844, 8.5845, 6.9522, 5.3848\}$$

(6) 得到相对残差序列

$$\varepsilon^{(0)} = \{-12.9831, 22.7892, -33.9394, -2.8437, 21.9588, -15.8705, -7.6955\}$$

(7) 状态划分。

状态的上、下限取为

$$\begin{cases} \Theta_{\min} = \min(\varepsilon^{(0)}) - 0.05 \times |\max(\varepsilon^{(0)}) - \min(\varepsilon^{(0)})| \\ \Theta_{\max} = \max(\varepsilon^{(0)}) + 0.05 \times |\max(\varepsilon^{(0)}) - \min(\varepsilon^{(0)})| \end{cases}$$

即残差所在区间上下各拓展 5%，算得 $\Theta = (-36.7758, 25.6256)$ 。

状态各小区间的间距取等间隔，状态个数则从 2 开始依次递增选取进行计算，最大值为序列长度。由此可得到 7 组试算结果，其中有以下情况之一的，认为该状态个数不可取：

情况一：某个状态小区间没有数据落入，即计算 P 的某行元素时分母为 0。

情况二：用 P 计算某状态的后续可能状态时，出现无法判断情况，而且向前后退至第一个数（多步转移）依然无法判断所求后续状态。

(8) 计算一步状态转移概率矩阵。

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0.08 & 0.42 & 0 & 0.5 \\ 0 & 0 & 0.3249 & 0.6751 \\ 0.5 & 0.5 & 0 & 0 \end{bmatrix}$$

(9) 使用状态转移概率矩阵得出预测结果，如表 6-13 所列。

表 6-13 预测结果

原数据	模型计算数据	原数据	模型计算数据
20.0000	20.0000	10.0000	10.5184
14.0000	15.8176	11.0000	10.4467
18.0000	16.9127	6.0000	6.1321
9.0000	10.6324	5.0000	6.5528
预测三个数	3.4219	2.4896	1.2732

(10) 计算预测精度。

首先，计算不同状态划分时的模型平均误差，如表 6-14 所列。可见，当状态个数取为 4 时，模型的精度最高。

其次，使用其他方法进行预测，并进行预测精度的比较，如表 6-15 所列。

表 6-14 不同状态划分下的模型平均误差

状态个数	模型平均误差	状态个数	模型平均误差
2	15.9536%	6	不可取
3	13.7272%	7	不可取
4	11.5191%	8	不可取
5	不可取	—	—

表 6-15 使用不同模型得到的预测精度

GM(1,1)模型	状态个数	经典模型	改进模型
16.8686%	2	15.9536%	15.9536%
	3	13.7272%	13.7272%
	4	不可取	11.5191%
	5	不可取	不可取
	6	不可取	不可取

当状态划分个数为 2、3 时，改进后的模型较 GM(1,1)模型的精度提高约为 1.06 倍和 1.23 倍，与经典的灰色-马尔可夫预测模型预测精度相同。当状态划分个数为 4 时，改进后的模型较 GM(1,1)模型的精度提高约为 1.46 倍，而经典的灰色-马尔可夫预测模型由于不能构建成功马尔可夫链状态转移概率矩阵，已不能取得预测结果。

可见随着状态划分个数的增加，改进模型的平均精度呈上升趋势，即改进模型支持更多的可划分状态个数，所以优于经典的灰色-马尔可夫预测模型。

灰色预测法应用环境相对宽松，不需要确定数据变动是否服从同一类型分布，不需要大的样本统计量就可以进行预测，具有一定的应用前景。

第 7 章

基于人工智能的武器装备备件 需求预测方法



智能化是预测技术发展的重要方向，机器学习则是提高预测智能化水平的主要途径。针对武器装备备件需求预测面临的数据样本少、建模难度大、不确定性强等问题及典型传统预测方法存在的不足，本章介绍武器装备备件需求智能预测方法。

7.1 基于 BP 神经网络的备件需求预测方法

武器装备备件种类繁多，影响其消耗的因素多，影响程度各异，导致在数据处理上需耗费大量精力。如果某些影响因素对预测结果造成的影响微小，就可以通过忽略这些因素来降低预测的难度，提高预测的速度，进而节约时间。反向传播（BP）网络结构简单、工作状态稳定，具有在任意精度内逼近非线性映射的能力，能够自动地逼近那些最佳刻画样本数据规律的函数，而且能处理数据中丢掉线性部分后残差的非线性。通过使用一个或多个隐含层，网络能够自动分割样本空间，并在不同的子空间里建立不同的函数，其灵活有效的学习方式、高度并行的处理机制使得 BP 算法能递进地接收数据进行学习并相应调整网络参数。此外，BP 网络还具有模拟多变量而不需要对输入变量做复杂的相关假定的能力。

它不依靠专家经验，只利用观察到的数据，可以从训练过程中通过学习抽取和逼近隐含的输入和输出的非线性关系。这对于建立基于多影响因素的备件需求量预测非常适用，因为很多影响因素与需求量间的关系都没有办法量化。

7.1.1 BP 神经网络基本原理

神经网络是在现代神经科学的基础上提出和发展起来的，旨在反映人脑结构及功能的一种抽象数学模型。自 1943 年美国心理学家 W.McCulloch 和数学家 W.Pitts 提出形式神经元的抽象数学模型——MP 模型以来，人工神经网络理论技术已经过了 50 多年曲折的发展，特别是 20 世纪 80 年代，人工神经网络的研究取得了重大进展，有关的理论和方法已经发展成一门介于物理学、数学、计算机科学和神经生物学之间的交叉学科。它在模式识别、图像处理、智能控制、组合优化、金融预测与管理、通信、机器人及专家系统等领域得到了广泛应用。学者共提出了 40 多种神经网络模型，其中比较著名的有感知器、Hopfield 网络、Boltzman 机、自适应共振理论及 BP 网络等。

误差反向传播网络是一种应用最广泛的神经网络。BP 网络模型是一个多层感知器，是由输入层、中间层（隐含层）和输出层构成的前馈网络，如图 7-1 所示。

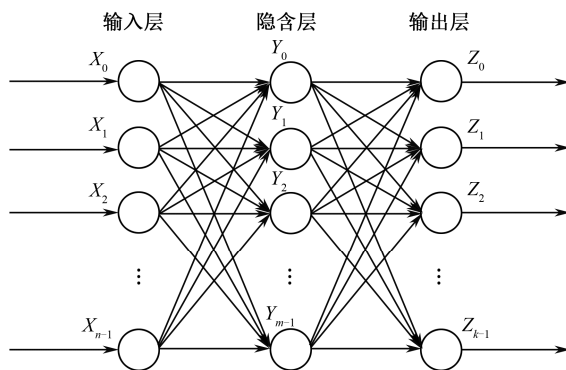


图 7-1 BP 网络模型

图 7-1 中有 n 个输入节点和 k 个输出节点，它们的非线性激活函数即 Sigmoid 函数有两种形式：

逻辑函数为

$$f(u) = \frac{1}{1 + e^{-u}} \quad (7-1)$$

双曲正切函数为

$$f(u) = \frac{1 - e^{-u}}{1 + e^{-u}} \quad (7-2)$$

隐含层的输出按下式计算：

$$Y_i = f\left(\sum_{j=0}^{n-1} W_{ij} X_j - \theta_i\right) \quad (i = 0, 1, \dots, m-1) \quad (7-3)$$

式中： X_i 为输入层第 i 个节点的输出值； W_{ij} 为输入层第 i 个节点到中间层第 j 个节点的权系数； θ_j 为中间层第 j 个节点的内部阈值。

输出层的输出按下式计算：

$$Z_l = f\left(\sum_{j=0}^{m-1} W_{jl} Y_j - \theta_l\right) \quad (l = 0, 1, \dots, k-1) \quad (7-4)$$

式中： Y_j 为中间层（隐含层）第 j 个节点的输出值； W_{jl} 为隐含层第 j 个节点到输出层第 l 个节点的权系数； θ_l 为输出层第 l 个节点的内部阈值； Z_l 为输出层中第 l 个节点的实际输出值。

BP 神经网络的自学习过程是一个反复迭代的过程，首先给网络一组初始权值，然后输入一个样本并计算其输出，通过实际的输出与期望值之间的差值用一定的方法来修改网络的权值，以达到减小这个差值的目的。反复执行这个过程直到这个差值小于预先确定的值为止。对足够的样本进行这样的训练后，网络所得的那组权值便是网络经过自适应学习得到的正确的内部关系。

BP 神经网络自适应学习过程的具体算法如下：

- (1) 给网络赋予一组小的随机初始权值，其值在 0~1，并使其互不相等。
- (2) 将输入数据规一化，使其在 0~1，根据所研究的问题确定期望输出信号 $(d_0, d_1, \dots, d_{k-1})$ 。
- (3) 逐层计算神经网络的实际输出值。
- (4) 从输出层开始，反向调整权值，其调整公式如下：

$$W_{jl} + \eta \delta_l Y_j \rightarrow W_{jl} \quad (7-5)$$

$$W_{ij} + \eta \delta_j X_i \rightarrow W_{ij} \quad (7-6)$$

式中

$$\delta_l = (d_l - Z_l) Z_l (1 - Z_l) \quad (7-7)$$

$$\delta_j = Y_j (1 - Y_j) \sum_{k=0}^{k-1} \delta_l W_{jl} \quad (7-8)$$

(5) 计算总的误差 E ，若 $E \leq \varepsilon$ ，则学习停止；否则，转到步骤(1)重新计算。

在实际编程序时，如果学习速率 η 较小，则学习速度较慢；而若 η 过大，则会引起网络出现摆动，所以采用学习速率可以变化的训练函数 `traingdx`。

Nielson (1989) 证明了对任意闭合区间连续函数都可以用含有一个隐含层的 BP 网络来逼近，因而 BP 网络可以用来预测备件的消耗量。

7.1.2 预测模型的建立

以影响备件需求量的因素作为输入，并以需求量作为输出，建立基于影响因素的 BP 网络预测模型，其基本步骤(图 7-2)如下：

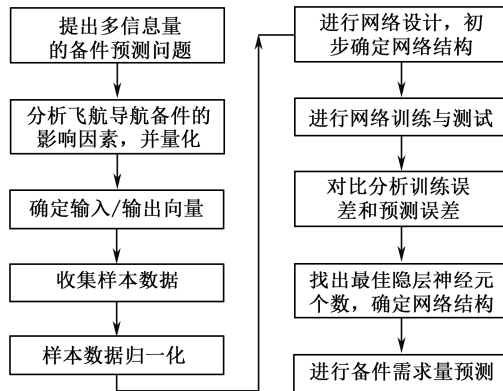


图 7-2 基于影响因素的 BP 网络预测流程

(1) 提出多信息量的备件预测问题，构思建立基于影响因素的 BP 网络预测模型。

(2) 分析并量化影响备件需求量的各因素。影响备件需求量的因素很多，如备件的固有故障率、重大演习任务、环境因素、使用维修与定检等。根据具体情况把各影响因素量化，并用适当的数字表示出来。

(3) 确定输入/输出向量，把影响因素量化值作为输入向量，把备件需求量作为输出向量。

(4) 收集样本数据并将其归一化。

(5) 确定输入与输出层神经元个数，并设计中间层神经元个数可变的网络结构。

(6) 输入训练样本与测试样本进行网络训练与测试, 对比训练误差和预测误差, 确定最优网络结构。

(7) 输入新的影响因素样本数据到确定的最优 BP 网络模型中进行预测。

7.1.3 示例分析

以某飞行团某型机载导弹的某备件为预测对象, 选取以下便于量化、易于统计的几个主要影响因素进行具体阐述:

年度重大演习任务次数(A): 一般重大演习任务都具有规模大、时效性强、使用的导弹装备和参与保障的人员多的特点, 在演习中难免会因装备使用强度高或精神紧张造成操作失误而导致备件需求。毫无疑问, 重大演习任务次数越多, 就越容易产生备件需求。

年度使用、维修与定检中该备件(或者含该备件的部件)被拆卸的总次数(B): 一般来说, 装备中某个部件被拆卸的次数越多, 该部件也相应容易出现故障。这是因为有的部件对于损坏的敏感性较高, 在装配、维修及使用过程中, 零部件有由于非正常因素而受到损坏的可能, 主要包括人为差错或使用不慎等, 例如, 备件因安装工具造成损坏。当对该件本身或在其附近对其功能有关的部件进行维修时, 也可以发生损坏。

年度累计挂机飞行时间(C): 导弹挂机飞行时处于高速、强振动的复杂运动状况, 对导弹自身的性能以及备件的性能都会有很大影响。挂机飞行时间越长, 次数越多, 备件需求量会越大。根据平时的训练计划可以统计出该型号导弹年度挂机飞行的总时间, 单位为小时。

使用与维修操作不当等人为因素损坏(D): 部队人员因退伍或转业会产生一定的流动性, 新上岗的操作与维修人员的个人素质、熟练程度及业务水平都需要一段时间的锻炼和培养, 这一过程中因操作不当及平时的疏忽等人为因素都会产生备件需求。此处以导弹技术准备过程中由于人为因素导致的备件失效数为该影响因素的量化值。

因自身固有故障率或到寿而自然消耗的备件个数(E): 备件的故障率是一种固有特性, 是其本身的设计和制造水平的体现。某备件在库存阶段因固有故障率引起的失效数量是可以统计出来的, 它直接影响着备件需求量。若此类消耗量较大, 则可以考虑选用替代品作为备件或者在订购环节重新选择生产商, 以提高备件的质量。把导弹装备定检过程中发现的失效备件数作为自然消耗的备件数。

环境因素(F): 高温、高湿、高盐、强日照、雨露、振动和灾害性气候等都能使导弹在使用与库存过程中产生备件需求。例如: 金属备件因腐蚀而导致强度降低、活动性变差、导电性减弱、断裂或破坏、活动体不能活动和电路断路等; 电子备件因受潮、高温、霉腐而导致参数漂移、动作失常、触点接触不良; 非金属备件的老化和变质等。而我国海军大部分岸舰、舰舰、潜舰导弹部署地都在沿海, 环境较为恶劣, 环境条件成为影响备件需求量的重要因素之一。根据该部队气象部门的统计资料, 把环境的严酷度分为7个等级, 并相应地量化为1~7范围的整数值, 数值越大, 表示严酷程度越高, 环境越恶劣。

1. 数据归一化处理

通过调研得到6个影响因素的相关数据及历史消耗数据, 如表7-1所列。

表7-1 影响因素量化值

年份	影响因素量化值						实际消耗量
	A	B	C	D	E	F	
2000	3	4	2	1	620	35	20
2001	2	3	4	2	840	30	15
2002	0	2	5	0	750	26	17
2003	1	2	2	1	680	25	25
2004	2	1	1	1	700	20	28
2005	3	5	3	2	560	30	32
2006	0	3	2	1	460	24	22
2007	1	2	1	1	580	20	26

在获得样本数据后, 由于其中各个指标因素互不相同, 原始样本中各量的单位与量级差别也不同, 为了计算方便及防止部分神经元达到过饱和状态, 需要对输入和输出向量进行归一化处理。利用 Matlab 实现样本数据的归一化过程, 将数据处理为区间[0,1]的数据。归一化方法很多, 本节采用如下公式:

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (7-9)$$

归一化后的样本数据见表7-2。

表 7-2 影响因素量化值

年份	影响因素量化值						实际 消耗量
	A	B	C	D	E	F	
2000	0.0000	0.0000	0.0000	1.0000	0.0000	1.0000	0.8182
2001	0.0000	0.1786	0.1667	0.8000	0.0000	0.6000	0.6818
2002	0.3333	0.2857	0.1852	0.6000	0.3333	0.4000	0.5909
2003	0.3333	0.3929	0.4259	0.2000	0.3333	0.2000	0.1818
2004	0.6667	0.5714	0.7222	0.2000	0.6667	0.4000	0.3182
2005	0.6667	0.7143	0.7778	0.4000	1.0000	0.4000	0.5000
2006	1.0000	1.0000	1.0000	0.4000	0.3333	0.8000	1.0000
2007	0.3333	0.3571	0.4074	0.0000	0.0000	0.0000	0.0000

2. 网络结构设计

由于单隐层 BP 网络的非线性映射能力很强，故采用单隐层的神经网络。由于输入样本为 6 维的输入向量，则输入层有 6 个神经元；而网络只有一个输出数据，那么输出层只有 1 个神经元。中间层的神经元个数需要通过实验来确定，为此需要设计一个隐含层可变的 BP 网络。一般可先根据 Kolmogorov 定理取 $2n+1$ 个，再依据经验在附近选择不同的个数进行试验寻找最佳数目。依实际情况分别选取 8、13、18、23、28 和 33 进行试验。按 BP 网络的一般设计原则，中间层神经元的传递函数为 S 型正切函数 `tansig`，输出层神经元的传递函数可设定为 S 型对数函数 `logsig`，这是因为输出已被归一化到区间[0,1]了，满足网络输出的要求。利用以下代码创建一个符合要求的 BP 网络：

```
a=[8 13 18 23 28 33];
for i=1:6
threshold=[0 1;0 1;0 1;0 1;0 1;0 1];
net=newff(threshold,[a(i),1],{'tansig','logsig'},'traingdx');
```

其中，`threshold` 设定了网络输入向量的取值范围[0,1]，网络所用的训练函数为 `traingdx`。该函数的学习算法是梯度下降动量法，并且学习速率是自适应的。

3. 最佳网络结构的确定

利用表 7-2 中 2001—2005 年的影响因素数据作为输入向量，以 2001—2006 年的消耗数量作为目标向量进行网络训练。训练参数的设定如表 7-3 所列，其他参数取默认值。需要说明的是，中间层神经元个数很难确定，但在很大程度上影

响着网络的预测性能。依照上面设计的网络，观察不同个数的隐单元构成的 BP 网络的训练曲线，分别检查网络性能，一般收敛速度较快且能够达到训练精度的网络为优。

网络训练代码如下：

```
net.trainParam.epochs=1000;
net.trainParam.goal=0.0001;
net=train(net,p,t);
```

表 7-3 训练参数

训练次数	训练目标
1000	0.0001

取 2006 年的影响因素数据作为测试样本，得到各网络的输出值后，分别与 2007 年的实际消耗量相比较，将此误差也作为选择最优网络结构的重要依据之一。一般取收敛速度最快且误差最小的网络结构为最优网络结构。

4. 预测

输入 2007 年的影响因素数据到最优网络模型中，预测 2008 年的备件需求量。

利用 Matlab 语言编制描述该 BP 网络预测模型的计算机程序，运行该程序可得到不同隐单元数的网络训练误差曲线，如图 7-3 所示。

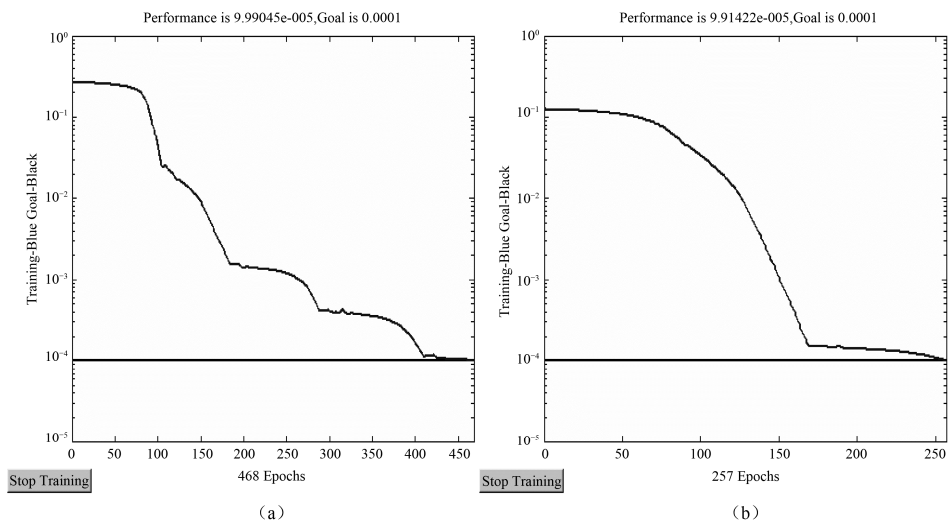


图 7-3 网络训练误差曲线对比

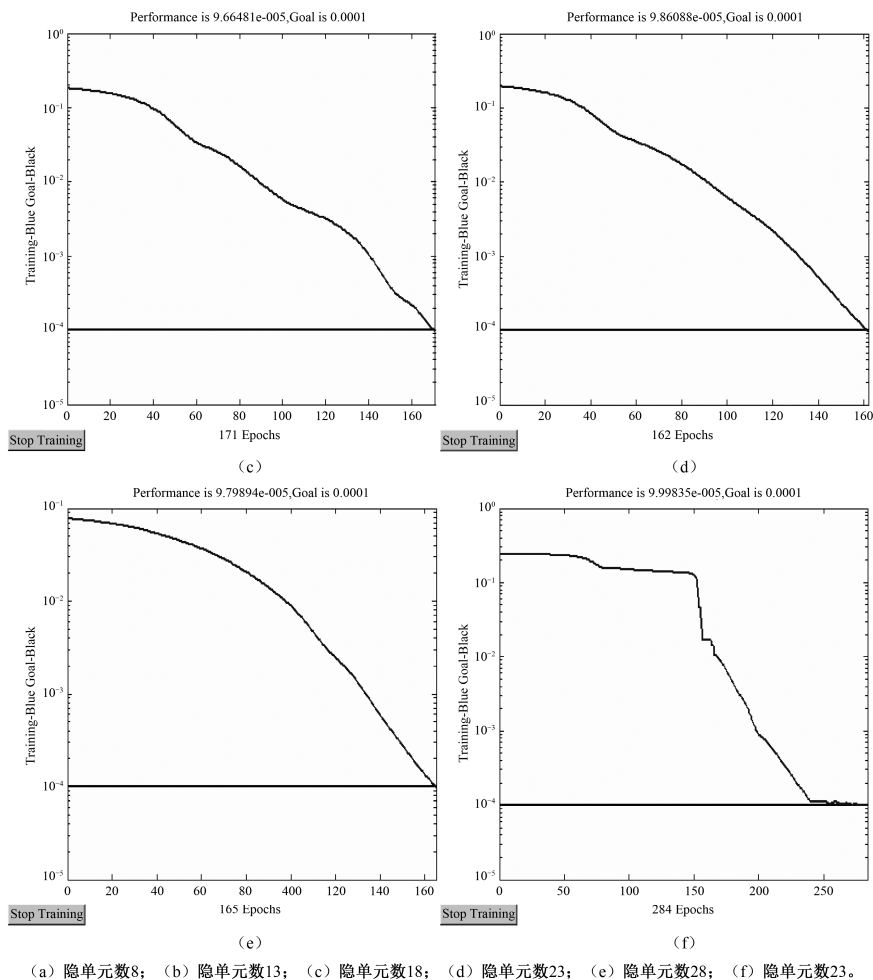


图 7-3 网络训练误差曲线对比 (续)

当各网络的训练精度都能达到时,为选择最优网络结构,需进行网络性能比较。从表 7-4 可以看出,训练次数与预测误差不随隐单元数的增加而线性增加或者减少,也证明了隐单元数的增加可以提高网络的映射精度,但并不意味着一定提高网络的性能的重要结论。很显然,当隐单元数为 23 时,网络只需要训练 162 次就能达到精度要求且测试数据的网络预测误差为 0.0456,在样本量比较少的情況下得到这样的效果说明该网络具备比较强的预测能力。

表 7-4 网络性能比较

隐单元个数	8	13	18	23	28	33
训练次数	468	257	171	162	165	284
测试数据的预测误差	0.1499	0.5755	0.2174	0.0456	0.6531	0.7211

以 2007 年的影响因素数据为输入, 网络的训练误差曲线如图 7-4 所示, 预测输出为 0.1686, 反归一化并取整数得 2008 年该备件的需求量为 75 个。

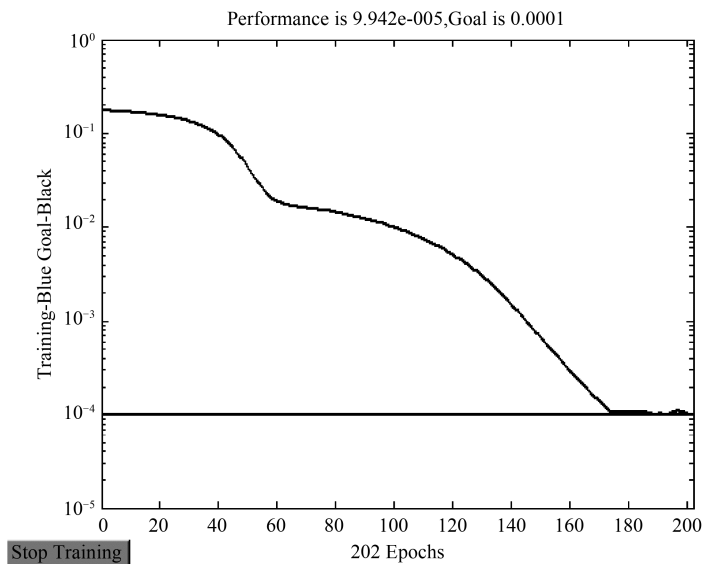


图 7-4 实际预测网络的训练误差曲线

7.2 基于支持向量机 (SVM) 的备件需求预测方法

传统的回归分析法需要有大量数据作为支撑, 而备件的消耗是一个复杂的非线性系统问题, 它受使用强度、使用的环境条件等因素影响, 在众多因素中, 有些因素可以量化, 而有些因素的作用机制却无法用精确的数学语言来描述。因此, 运用传统的回归分析法进行备件的消耗预测存在较大的误差。

7.2.1 支持向量机概述

回归分析法是当前预测领域中应用较为广泛的一种方法,不过由于该方法需要大量的历史数据,并且利用回归分析法所得到的回归预测方程往往只能考虑少数几种影响因素,而略去了好多未考虑的因素,在一定程度上影响了预测的准确性。基于支持向量回归的预测方法,通过实践运用,证实为解决复杂多因素预测问题时有较大优势。

支持向量机(Support Vector Machines, SVM)是 20 世纪 90 年代中期由 Vapnik 教授领导的研究小组提出的一种新的智能机器,源于 20 世纪 70 年代迅速发展起来的统计学习理论,由于具有较完备的理论基础和较好的学习性能,能很好地解决小样本、非线性、高维数和局部极小点等实际问题,尤其对小样本问题有非常好的拟合效果,现已广泛运用到模式识别、函数逼近、回归估计等领域。

支持向量机大体上分为支持向量机分类(SVC)和支持向量机回归(SVR)两个方面。SVC 用来处理分类问题,既可以处理两类分类问题,又可以处理多类分类问题;SVR 则用来处理函数回归问题。

回归问题就是给定一个新的输入样本 \mathbf{x} , 根据给定的数据样本推断它所对应的输出 y 是多少,这个输出 y 是一个实数,用数学语言可以把回归问题描述如下:

给定的数据样本集合为 $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\}$ 。其中 $\mathbf{x}_i \in \mathbf{R}^n$, $y_i \in \mathbf{R}$, $i=1, 2, \dots, l$ 。寻找 \mathbf{R}^n 上的一个函数 $f(\mathbf{x})$, 以便用 $y = f(\mathbf{x})$ 来推断任一输入 \mathbf{x} 所对应的 y 值。

分类问题和回归问题的结构相同,不同之处仅在于它们的输出 y_i 和 y 取值范围不同。在分类问题中,输出只允许取 -1 或 1 两个值或有限个值;在回归问题中,输出可取任意实数。

支持向量机的主要优点如下:

(1) 它是专门针对有限样本情况的,其目标是得到现有信息下的最优解而不仅仅是样本数趋于无穷大时的最优解。

(2) 算法最终将转化为一个二次型寻优问题,从理论上说,得到的将是全局最优解,解决了在神经网络方法中无法避免的局部极值问题。

(3) 算法将实际问题通过非线性变换到高维的特征空间,在高维空间中构造线性判别函数来实现原空间中的非线性判别函数,特殊性质能保证机器有较好的推广能力,同时它巧妙地解决了维数问题,其算法复杂性与样本维数无关。

7.2.2 支持向量机预测原理

1. 支持向量机回归问题

SVM 回归问题与分类问题有一些类似, 给定的数据样本集合为 $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\}$, 其中 $\mathbf{x}_i \in \mathbf{R}^n$, $y_i \in \mathbf{R}$ ($i=1, 2, \dots, l$)。与分类问题不同, 这里的 y_i 并不限定取 -1 或 1 , 而是可取任意实数。

回归问题就是给定一个新的输入样本 \mathbf{x} , 根据给定的数据样本推断它所对应的输出 y 是多少, 这个输出 y 是一个实数, 用数学语言可以把回归问题描述如下:

给定的数据样本集合为 $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\}$, 其中 $\mathbf{x}_i \in \mathbf{R}^n$, $y_i \in \mathbf{R}$ ($i=1, 2, \dots, l$)。寻找 \mathbf{R}^n 上的一个函数 $f(\mathbf{x})$, 以使用 $y = f(\mathbf{x})$ 来推断任一输入 \mathbf{x} 所对应的 y 值。

由此可见, 回归问题和分类问题的结构相同, 不同之处仅在于它们的输出 y_i 和 y 取值范围不同。在分类问题中, 输出只允许取两个值或有限个值; 在回归问题中, 输出可取任意实数。

当 $n=1$ 时, 上述回归问题有着明显的几何意义, 如图 7-5 所示。

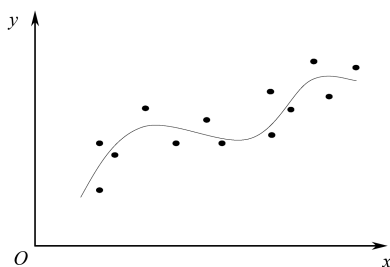


图 7-5 回归问题的几何意义

图 7-5 中, “ \cdot ” 表示给定的数据样本集中的样本点, 回归所要寻找的函数 $f(\mathbf{x})$ 对应图中的曲线。

支持向量的方法应用到回归问题中, 保留了最大间隔算法的所有的主要特征, 非线性函数可以通过核特征空间中的线性学习器得到。与 SVM 分类算法一样, SVM 回归算法要最小化一个凸函数, 并且它的解是稀疏的。

与分类算法的思路一样, 回归算法需要定义一个损失函数, 该函数可以忽略真实值某个上下范围内的误差, 这种类型的函数也就是 ϵ 不敏感损失函数。

图 7-6 显示了具有 ε 不敏感带的一维线性回归函数的例子。变量 ε 度量了训练点上误差的代价，在 ε 不敏感带内点误差为 0。

图 7-7 显示了非线性回归函数的不敏感带。损失函数的解以函数的最上化为特征，使 ε 不敏感损失函数就有这个优势，以确保全局最小解的存在和可靠泛化界的优化。

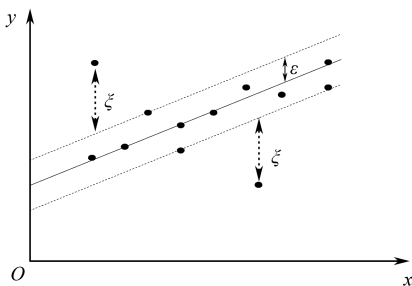


图 7-6 一维线性回归函数的不敏感带

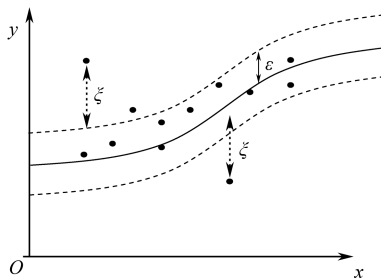


图 7-7 非线性回归函数的不敏感带

2. 线性支持向量机回归

样本数据集为线性时，设数据样本为 n 维向量，某区域的 l 个数据样本及其值的表示为 $(\mathbf{x}_i, y_i) (i=1, 2, \dots, l, \mathbf{x} \in \mathbf{R}^n, y \in \mathbf{R})$ 。

线性函数设为

$$f(\mathbf{x}) = \boldsymbol{\omega} \cdot \mathbf{x} + b \quad (7-10)$$

优化问题即最小化：

$$R(\boldsymbol{\omega}, \xi, \xi^*) = \frac{1}{2} \boldsymbol{\omega} \cdot \boldsymbol{\omega} + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (7-11)$$

约束条件为

$$\begin{cases} y_i - f(\mathbf{x}_i) \leq \xi_i^* + \varepsilon \\ f(\mathbf{x}_i) - y_i \leq \xi_i + \varepsilon \quad (i=1, \dots, l) \\ \xi_i^*, \xi_i \geq 0 \end{cases}$$

式中： $C > 0$ 为惩罚系数， C 越大表示对超出 ε 管道数据点的惩罚越大。

式 (7-11) 中：第一项使函数更为平坦，从而提高了泛化能力；第二项减小误差，常数 C 对两者做出折中。 ε 为一正数， $f(\mathbf{x}_i)$ 与 y_i 的差别小于 ε 时不计入误差，大于 ε 时误差计为 $|f(\mathbf{x}_i) - y_i| - \varepsilon$ 。

这也是一个凸二次优化问题，引入拉格朗日函数：

$$\begin{aligned}
 & L(\boldsymbol{\omega}, b, \xi, \xi^*, \alpha, \alpha^*, \gamma, \gamma^*) \\
 &= \frac{1}{2}(\boldsymbol{\omega} \cdot \boldsymbol{\omega}) + C \sum_{i=1}^l (\xi_i + \xi_i^*) - \sum_{i=1}^l \alpha_i [\xi_i + \varepsilon + y_i - f(\mathbf{x}_i)] \\
 & \quad - \sum_{i=1}^l \alpha_i^* [\xi_i^* + \varepsilon + f(\mathbf{x}_i) - y_i] - \sum_{i=1}^l (\xi_i \gamma_i + \xi_i^* \gamma_i^*)
 \end{aligned} \tag{7-12}$$

式中: $\alpha_i, \alpha_i^* \geq 0; \gamma_i, \gamma_i^* \geq 0; i=1, 2, \dots, l$ 为拉格朗日乘子。

因此

$$\frac{\partial L}{\partial \boldsymbol{\omega}} = 0, \quad \frac{\partial L}{\partial b} = 0, \quad \frac{\partial L}{\partial \xi_i} = 0, \quad \frac{\partial L}{\partial \xi_i^*} = 0 \tag{7-13}$$

从而有

$$\sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \tag{7-14}$$

$$\boldsymbol{\omega} = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \mathbf{x}_i \tag{7-15}$$

$$C - \alpha_i - \gamma_i = 0 \tag{7-16}$$

$$C - \alpha_i^* - \gamma_i^* = 0 \tag{7-17}$$

将式(7-14)~式(7-17)代入式(7-12)中, 所要求解的最小值问题可以转化为求解对偶问题的最大值问题:

$$\begin{aligned}
 W(\alpha, \alpha^*) = & -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(\mathbf{x}_i \cdot \mathbf{x}_j) + \sum_{i=1}^n (\alpha_i - \alpha_i^*) y_i \\
 & - \sum_{i=1}^n (\alpha_i + \alpha_i^*) \varepsilon
 \end{aligned} \tag{7-18}$$

约束条件为

$$\begin{cases} \boldsymbol{\omega} = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \mathbf{x}_i \\ 0 \leq \alpha_i \leq C (i=1, 2, \dots, l) \\ 0 \leq \alpha_i^* \leq C (i=1, 2, \dots, l) \end{cases} \tag{7-19}$$

这也是一个二次优化问题, $\boldsymbol{\omega}$ 可由式(7-13)得到, b 的求法与分类情况相同。

3. 非线性支持向量机回归

对于非线性训练集, 可以通过非线性函数 $\phi(x)$ 将训练集数据 \mathbf{x} 映射到一个高

维线性特征空间, 在这个维数可能为无穷大的线性空间中构造回归估计函数, 则估计函数可表示为

$$f(\mathbf{x}) = \boldsymbol{\omega}\phi(\mathbf{x}) + b \quad (7-20)$$

式中: $\boldsymbol{\omega}$ 的维数为特征空间维数。

其关键问题也是核函数 $K(\mathbf{x}, y)$ 的使用。这里, 优化问题成为在式 (7-19) 的约束下最大化函数

$$\begin{aligned} W(\alpha, \alpha^*) = & -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(\mathbf{x}_i, \mathbf{x}_j) + \sum_{i=1}^n (\alpha_i - \alpha_i^*) y_i \\ & - \sum_{i=1}^n (\alpha_i + \alpha_i^*) \varepsilon \end{aligned} \quad (7-21)$$

此时有

$$\boldsymbol{\omega} = \sum_{i=1}^l (\alpha_i - \alpha_i^*) \phi(\mathbf{x}_i) \quad (7-22)$$

函数 $f(\mathbf{x})$ 可直接表示为

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(\mathbf{x}_i \cdot \mathbf{x}) + b \quad (7-23)$$

按照 Kuhn-Tucker 定理, 有

$$\alpha_i [\varepsilon + \zeta_i - y_i + f(\mathbf{x}_i)] = 0 \quad (i=1, 2, \dots, l) \quad (7-24)$$

$$\alpha_i^* [\varepsilon + \zeta_i^* - y_i + f(\mathbf{x}_i)] = 0 \quad (i=1, 2, \dots, l) \quad (7-25)$$

$$\xi_i \gamma_i = 0 \quad (i=1, 2, \dots, l) \quad (7-26)$$

$$\xi_i^* \gamma_i^* = 0 \quad (i=1, 2, \dots, l) \quad (7-27)$$

由式 (7-24) 和式 (7-25) 可见, $\alpha_i \alpha_i^* = 0$, 即任何一组 α_i 和 α_i^* 都不会同时为非 0。由式 (7-14) ~ 式 (7-17) 可得

$$(C - \alpha_i) \xi_i = 0 \quad (i=1, 2, \dots, l) \quad (7-28)$$

$$(C - \alpha_i^*) \xi_i^* = 0 \quad (i=1, 2, \dots, l) \quad (7-29)$$

由此可见, 对应于 $\alpha_i = C$ 或 $\alpha_i^* = C$ 的 $f(\mathbf{x}_i)$ 与 y_i 的误差可能大于 ε , 对应 $\alpha_i \in (0, C)$ 或 $\alpha_i^* \in (0, C)$ 的 $f(\mathbf{x}_i)$ 与 y_i 的误差必然等于 ε , 即 $\xi_i = 0$ 或 $\xi_i^* = 0$ 。

因此有

$$\varepsilon - y_i + f(\mathbf{x}_i) = 0 \quad (\alpha_i \in (0, C)) \quad (7-30)$$

$$\varepsilon - y_i + f(\mathbf{x}_i) = 0 \quad (\alpha_i^* \in (0, C)) \quad (7-31)$$

由式 (7-23)、式 (7-30) 和式 (7-31) 可以求出 b 。

4. 核函数的类型选择

SVM 利用核函数实现输入数据到高维特征空间的映射, 从而实现算法从线性到非线性情况下的推广。核函数表述了学习算法搜索的假设空间的结构和特点, 核函数的选择隐含地确定了所映射的特征空间, 核函数的形式及参数决定了学习机器的类型和复杂程度, 选择合适的核函数, 可以有效地提高核函数的性能。

选取合适的核函数可以提高精度。支持向量机常用的核函数有以下四种:

(1) 线性核函数, 即 $K(x^r \cdot y)$ 。

(2) 多项式核函数, $K(\mathbf{x}, \mathbf{x}_i) = |(\mathbf{x} \cdot \mathbf{x}_i) + 1|^q$, 对应 SVM 是一个 q 阶多项式分类器。

(3) 高斯径向基函数核函数 (Radial Basic Function, RBF):

$K(\mathbf{x}, \mathbf{x}_i) = \exp \left| -\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{g^2} \right|$, 对应 SVM 是一种径向基函数分类器。

(4) Sigmoid 核函数, $K(\mathbf{x}, \mathbf{x}_i) = \tanh(s(\mathbf{x} \cdot \mathbf{x}_i) + c)$ 。

在选择核函数时, 一般首选径向基核函数, 与其他核函数相比较, 其有以下优点:

(1) 径向基核函数将样本集非线性映射到一个高维空间, 与多项式核函数不同的是, 在样本特征是非线性情况时, 它能够处理这种非线性情况。而实际上多项式核函数是径向基核函数的一个特殊情况。另外, Sigmoid 核函数在某些特定参数情况下与径向基核函数相似。

(2) 考虑参数个数对模型选择复杂度的影响而选择径向基核函数。参数越多, 模型选择工作量就越大。多项式核函数和 Sigmoid 核函数需要确定的参数大于径向基核函数: 多项式核函数和 Sigmoid 核函数都需要确定两个核参数, 而径向基核函数只需确定一个核参数 g 。

(3) 径向基核函数容易进行数值计算。径向基核函数的值大于或等于 0、小于或等于 1 ($0 \leq K(\mathbf{x}, \mathbf{x}') \leq 1$), 而多项式核函数的值可能为无穷大或者零。Sigmoid 核函数在一定参数情况下并不一定总是正定的, 因此不一定满足 Mercer 条件。所以接下来的试验都选择用径向基核函数来构造支持向量机分类模型。

5. 支持向量机的参数优化

对于支持向量机而言, 在核函数选定的情况下, 影响其性能的关键因素是核函数的参数及误差惩罚系数 C , 为保证学习机器具有最佳推广能力, 需要进行参

数的优化。

核函数参数主要影响样本数据在高维特征空间中分布的复杂程度,核参数的改变实际上隐含着特征空间的 VC 维的改变,从而影响置信范围,最终影响结构风险范围。误差惩罚系数 C 控制间隔的最大化与分类误差之间的折中: C 越大,则对于错分样本的惩罚越大; C 的取值小则表示惩罚小。学习机器的复杂度小而经验风险值较大,前者称为“过学习”,而后者称为“欠学习”。除了在同一特征空间优化 C 以获得对应空间的最优 SVM,还要优化核函数参数以获得全局最优的 SVM。

由于径向基核函数的特点,常选择其作为支持向量机的核函数。当使用径向基核函数时,有两个重要参数 C 和 g 需要优化。选择合适的参数 C 和 g 能够使支持向量机的预测取得优秀的性能。就支持向量机理论目前的研究现状而言,还不能从理论上给出合适参数的选择方法,所以,只有通过试验的方式来进行参数的寻优。

相比于核函数,核参数是影响 SVM 性能的关键因素。国内外许多研究者针对 SVM 核参数的选择已提出许多算法,常见的有网格搜索法、拟牛顿法、遗传算法等。

1) 网格搜索法

网格搜索法是将 C 和 g 分别取 M 个值和 N 个值,对 $M \times N$ 个 (C, g) 的组合,分别训练不同的支持向量机,再估计其学习精度,从而在 $M \times N$ 个 (C, g) 的组合中得到学习精度最高的一个组合作为最优参数。网格搜索法的优点是可以并行处理,每个 SVM 的训练都是独立的;其缺点是计算量比较大。改进的网格搜索法其基本思想是:首先用较大步长的 (C, g) 组合,得到最优的 (C, g) 值;然后在 (C, g) 值旁的一定范围内进行一次更细致的网格搜索。

2) 拟牛顿法

拟牛顿法在 SVM 训练之后,使用一个 Sigmoid 函数将 SVM 在校验集上的输出映射为相应的概率。该概率是一个后验概率,用于校验集上的误差估计 $E(\alpha, \theta)$;然后通过最小化误差估计,来得到超参数的最优值,从而完成优化。

拟牛顿法的优化算法:

(1) 根据经验初始化 θ 为 θ_0 , 使得 $\theta_0 = \arg \min_{\theta} E(\alpha_0, \theta)$ 。

(2) 在训练集上,使用一个标准 SVM 算法,用 θ_0 训练 SVM,得到一个 SVM 模型。

(3) 在验证集上用牛顿法优化推断参数 A 和 B 的最佳值。

(4) 计算误差估计的梯度 $\frac{\partial}{\partial \theta} E(\alpha, \theta)$ 。

(5) 在内核参数空间上计算海赛矩阵 \mathbf{H}' 。

(6) 用 $\Delta \theta = -\lambda \mathbf{H}' \frac{\partial}{\partial \theta} E(\alpha, \theta)$ (其中: λ 为沿着搜索方向的步长, \mathbf{H}' 为 $n \times n$

矩阵, n 为向量 θ 的维数) 来更新 θ 直至误差估计 E 被最小化。

3) 遗传算法

遗传算法 (Genetic Algorithms, GA) 是由美国 Holland 教授在 1975 年提出的一种模拟达尔文遗传选择和生物进化过程的随机搜索算法。其主要特点是不依赖于梯度信息的隐含并行随机群体搜索, 能在搜索过程中自动获取和积累有关搜索空间知识, 并能自适应地控制搜索进程, 从而得到最优解。遗传算法从初始种群出发, 采用“优胜劣汰, 适者生存”的自然法则选择个体, 并通过交叉、变异来产生新一代种群, 如此逐代进化, 直到满足目标为止。遗传算法对于复杂的优化问题无须建模和复杂运算, 只要利用遗传算法的三种算子就能得到最优解。标准遗传算法的构成包括染色体的编码、种群规模、适应度函数、遗传算子。

(1) 染色体的编码。编码是指将问题的解空间转换成遗传算法所能处理的搜索空间。常用的编码方式有二进制编码、符号编码和浮点数编码。二进制编码、解码操作简单易行, 遗传操作便于实现, 被广泛应用。

(2) 种群规模。群体规模直接影响遗传算法的性能, 一般来说, 种群规模在 20~200 范围, 能够很好地实现种群多样性与算法复杂度之间的折中。

(3) 适应度函数。适应度函数是用于评价各码串对问题适应程度的准则, 它是遗传算法指导搜索的唯一信息, 它的好坏是衡量算法优劣的关键。

(4) 遗传算子。

①选择算子: 将父代中适应度值高 (性能优良) 的染色体 (个体) 选择复制到下一代中, 并淘汰劣势个体, 选择过程通常采用适应度比例选择法 (又称为赌轮选择法), 适应度越高的个体被复制到下一代的概率越高。

②交叉算子: 从群体中随机选择一对个体, 以交叉概率 p_c ($0 < p_c \leq 1$) 实施交叉操作, 即互相交换各自的部分基因 (表示染色体的数字串的某一位), 从而形成一对新的个体, 常见的交叉方式有单点交叉、两点交叉、多点交叉及均匀交叉等。交叉概率 p_c 一般取 0.6~1.0。

③变异算子: 以很小的概率随机地改变遗传基因的值, 在染色体以二进制编码的算法中, 随机地将染色体的某一位基因由 1 变为 0, 或由 0 变为 1。通过变异算子, 可确保群体中基因类型的多样性, 防止搜索停滞, 变异概率 p_m 一般取 0.001~0.1。

应用遗传算法对支持向量机进行参数优化的具体步骤如图 7-8 所示。

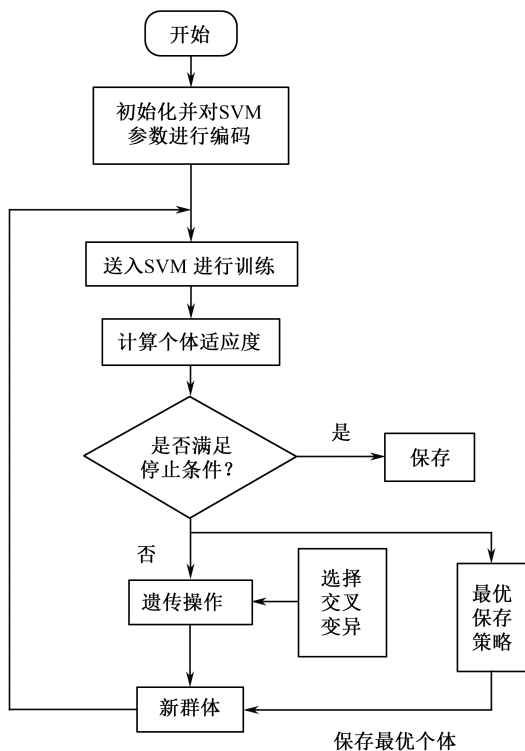


图 7-8 遗传算法优化支持向量机参数流程

(1) 初始化参数并对其进行二进制编码，产生遗传算法的初始群体；群体数据选取 100，二进制编码码串的长度根据要搜索的支持向量机参数的范围与精度来确定。当核函数选择 RBF 核函数时，支持向量机需要优化的参数为 C 与 g ，码串中包括两个参数的二进制码。

(2) 设定遗传算法的参数，初始代数设为 $G=1$ ，初始交叉概率 $p_c(1)=0.8$ ，初始变异概率 $p_m=0.008$ ，最大遗传代数 $G_{\max}=100$ 。

(3) 将初始群体解码后送入支持向量机进行训练并计算个体的适应度。当进行分类时，适应度函数采用支持向量机分类的正确率；当进行预测时，适应度函数采用支持向量机预测结果与实际结果的方差。

(4) 应用最优保存策略，在进行遗传操作之前先把适应度最高的个体保存下来，以防止优秀基因因遗传算子操作而丢失；记录最差个体的序号 Index。

(5) 进行遗传操作，选择算子采用赌轮选择法，交叉算子采用单点交叉，经

过遗传算子操作产生新群体，并用步骤（4）保存的最优个体替换序号为 Index 的新个体，并最终产生新的群体，然后返回步骤（3）进行训练。

（6）检查是否算法终止条件，由于适应度本身就是搜索的结果，很难作为终止条件，但是当连续几代最优个体的适应度相等时，则认为种群不能再进行优化，作为算法终止的条件之一，把设定的最大遗传代数也作为算法的终止条件之一。当满足上述两个条件中的任何一个时，则自动终止算法。

遗传算法是近年来才被提出并应用于 SVM 核参数的选择上的。相比于传统的迭代法，遗传算法对目标函数的要求不高，不容易陷入局部最优，能够在很大程度上提高 SVM 的分类精度及核参数的搜索速度。

6. 支持向量机训练、测试步骤

运用基于支持向量机的方法进行分类及预测过程中，都会涉及支持向量机训练、测试及核函数选择等重要问题。支持向量机训练和测试步骤如下：

- （1）将数据转换成支持向量机要求的格式，并对数据进行标准化处理。
- （2）将处理后的数据作为支持向量机训练机的输入，进行训练。
- （3）尝试使用不同核函数并比较结果，以选择一个较好的核函数。
- （4）对核函数参数及误差惩罚系数进行优化，以获得最优参数。
- （5）使用最优参数对整个训练集进行训练，得到支持向量机预测器。

（6）将测试数据集输入支持向量机预测器进行分类或预测，并对所得结果的各种参数进行评价。

7.2.3 支持向量机预测模型

1. 预测步骤

以影响备件需求的因素作为输入，以备件的需求量作为输出，建立基于支持向量机的备件需求预测模型。其基本步骤（图 7-9）如下：

- （1）分析影响备件需求的因素。
- （2）影响因素的量化处理。通过量化处理使需要定量数值描述的影响因素得到定量值，从而方便进行分析计算。
- （3）对影响因素量化后的数据进行规范化处理，方便各种影响因素之间进行分析比较。
- （4）采用支持向量机对备件需求量样本进行学习，备件需求的影响因素是输入，备件需求量是输出；并对精度进行分析，若不能达到要求的精度，转至步

骤（3），重新选择核函数，获得最优参数；若达到精度要求，生成合适的 SVM 预测函数。

（5）输入新的备件需求量样本或者备件在不同情况下的需求量样本，到支持向量机模型中进行备件需求预测。

（6）增添新的备件需求样本到支持向量机进行学习，不断地提高备件需求预测的精度。

（7）预测备件需求数据，保存模型。

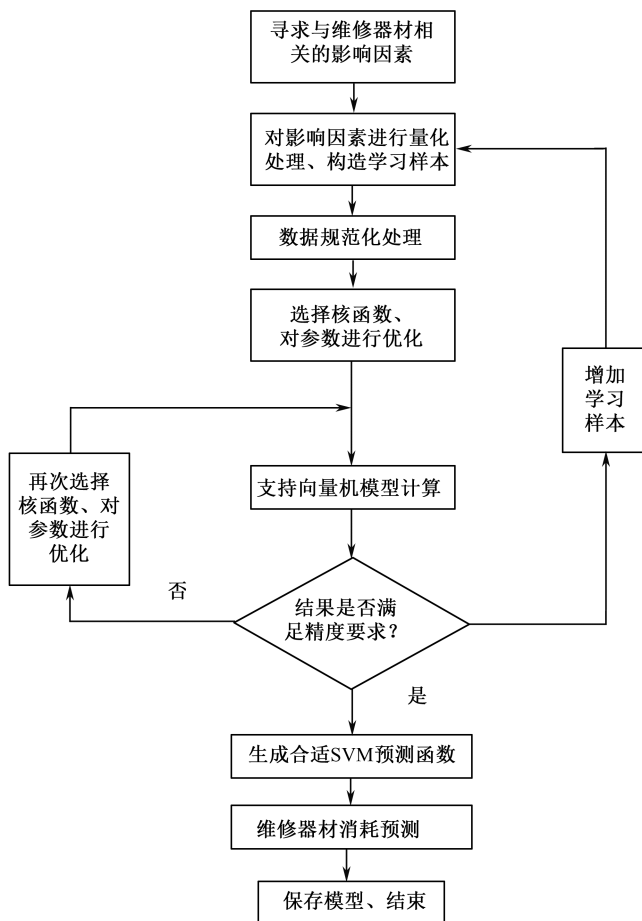


图 7-9 基于支持向量机的备件需求预测流程

2. 备件需求影响因素分析与量化

根据对某型飞机航材保障情况的分析研究，影响此机型航材需求的主要因素

包括飞行训练任务情况、机务人员的技术条件和航材综合性能。

(1) 飞行训练任务情况：飞行的时间、起落、科目等，飞行时间越长，起落越多，飞机发生故障的概率也就越大，航材备件需求量增加；飞行科目的不同使得各个机件的使用频率不同，越是频繁使用的航材发生故障的概率也就越大，如在起落训练中，轮胎和起落架发生故障的概率就高，相应航材备件的需求量增加。

(2) 机务人员的技术条件：外场机务人员的维护水平及修理厂对故障件的修复能力都对航材的消耗产生不同的影响。机务人员维护水平越高，修理厂对故障件的修复能力越强，航材备件的需求量就会越小。

(3) 航材综合性能：制造工艺、产品质量等也会对航材的消耗产生重要的影响。航材质量越好，平均故障间隔时间越小，航材备件的需求量越小。

(4) 航材的采购的难易程度：航材易于采购，采购时间短，则航材备件的储备量就不用太多；反之，航材采购困难，采购时间长，则储备量就要相对多一些，需求量也就更高一些。

根据此原则将航材备件需求量 L 的主要影响因素归纳为飞行时间 T 、飞行起落数 Y 、机务综合业务素质差的人员在全体机务人员所占的比例 U 、故障率 Q 和不易采购的程度 S 。将影响航材备件需求的五个因素作为支持向量机的输入参数，将航材备件需求量 L 作为输出。故基于支持向量机的航材需求量为

$$L = f(T, Y, U, Q, S)$$

在上述五个影响因素中，前四个指标均为数值指标，很容易量化。对于第五个因素，可以通过查询航材的订货记录结合航材工作人员的经验，将这一指标转化为 0~1 的数值指标，即容易采购则赋值为 1，否则为 0。根据基于影响因素的支持向量机航材需求预测流程，对航材需求进行预测。

3. 需求量预测准确度评价

对于航材备件需求量的预测准确度评，可以从常用的需求预测评价方法中选取合适的指标作为评价标准。选用平均绝对误差（Mean Absolute Error, MAE）、均方误差（Mean Squared Error, MSE）对需求量的预测结果进行综合评价：

$$\text{MAE} = \frac{\sum_{t=1}^n |X'_t - X_t|}{n}$$

$$\text{MSE} = \frac{\sum_{t=1}^n (X'_t - X_t)^2}{n}$$

根据上述定义可知，MAE、MSE 值越小，表明预测值与实际值的拟合精度越高，则需求预测的准确度越高。

7.2.4 示例分析

以某型飞机的航材库存管理实践为背景，进行数据试验与分析。该机型在海军装备数量很少，航材的消耗样本小。为了保证飞机的正常训练和特殊飞行，该团航材股对每一种航材都保持了一定的库存，但大部时间部分航材都处于闲置状态，这显然占用了大量资金，经济性不好。如何依据现有资料，寻求到航材的需求规律，预测出未来一段时间内航材的需求量情况，决定航材的合理储存量，进行库存结构的优化，是保障飞行、节约经费、减少航材积压的重要课题。

这里的仿真程序选用 Matlab 2008a 软件来实现，其中的支持向量机组件是由台湾大学林智仁教授编写的 Libsvm 工具箱，试验中的支持向量机采用高斯核函数。首先用指数平滑法对航材的需求数据进行预测，其次分别使用网格搜索法和遗传算法对支持向量机的核参数 g 和惩罚系数 C 进行优化，然后将两种方法优化参数的支持向量机预测模型对航材的历史需求数据进行需求预测，最后将三种方法得到的预测数据进行对比。

1. 数据准备

由于航材发展的同源性和继承性，飞行任务的相近性，同类航材的任务相近、工作原理相似，因此航材相似性是普遍存在的。在环境、使用条件、使用场合和管理水平等因素相近的情况下，相似航材具有相似的需求特征，同时考虑到数据试验的复杂程度，这里随机选取 A 类航材中具有代表性的航材历史数据来例证基于支持向量机的航材需求预测方法。

以年为单位，对冷气电磁阀 2000—2009 年需求影响因素的相关数据进行收集整理，结果见表 7-5。以前 9 年的数据作为训练样本，以第 10 年（2009 年）的数据作为测试样本，应用基于支持向量机的航材需求模型进行需求量预测。

表 7-5 航材各年需求影响因素的相关数据（冷气电磁阀）

年份	T/h	$Y/次$	$U/\%$	$Q/（个/10^3h）$	S	$L/个$
2000	5362	854	0.089	0.0034	0.35	5
2001	6548	921	0.102	0.0045	0.41	7
2002	6209	875	0.096	0.0041	0.33	6
2003	6859	958	0.129	0.0050	0.52	8

续表

年份	T/h	$Y/\text{次}$	$U/\%$	$Q/(\text{个}/10^3\text{h})$	S	$L/\text{个}$
2004	6925	1205	0.151	0.0063	0.69	10
2005	6891	1181	0.129	0.0059	0.62	9
2006	6722	973	0.093	0.0056	0.58	8
2007	6797	1025	0.142	0.0064	0.71	9
2008	7132	1266	0.168	0.0073	0.83	11
2009	8643	1526	0.193	0.0081	0.87	13

2. 数据预处理

为了适应支持向量机的对数据的要求,减少误差,进而更加准确地进行需求预测,采用极差变换法分别对选择的航材的历史需求数据按照下列步骤进行归一化处理:

- (1) 寻找历史需求数据中的最大值和最小值,分别记为 MAX, MIN;
- (2) 将第 t 年的需求量 X_t 转化为

$$X'_t = \frac{X_t - \text{MIN}}{\text{MAX} - \text{MIN}}$$

通过极差变换法,将航材的每年影响航材需求的相关数据均变为属于 $[0,1]$ 区间的值。编号为 01054021 的航材各年需求影响因素的相关数据经归一化处理的结果见表 7-6。极差变换归一化处理并不改变各期数据之间的相对关系,因此,不会影响需求预测模型的构建。

表 7-6 归一化航材各年需求影响因素的相关数据 (冷气电磁阀)

年份	T/h	$Y/\text{次}$	$U/\%$	$Q/(\text{个}/10^3\text{h})$	S	$L/\text{个}$
2000	0.620386	0.09877	-3×10^{-5}	-0.00058	4.05×10^{-5}	0.0005
2001	0.757607	0.10651	-3.6×10^{-5}	-0.00081	4.74×10^{-5}	0.0008
2002	0.718385	0.10120	-2.7×10^{-5}	-0.00069	3.82×10^{-5}	-0.9993
2003	0.79359	0.11079	-4.5×10^{-5}	-0.00093	6.02×10^{-5}	0.0009
2004	0.801226	0.13935	-6.2×10^{-5}	-0.00116	7.98×10^{-5}	0.0011
2005	0.797292	0.13658	-5.7×10^{-5}	-0.00104	7.17×10^{-5}	0.0010
2006	0.777739	0.11251	-5.6×10^{-5}	-0.00092	6.71×10^{-5}	0.000926
2007	0.786416	0.11852	-6.6×10^{-5}	-0.00104	8.21×10^{-5}	0.0010
2008	0.825176	0.14640	-7.7×10^{-5}	-0.00127	9.6×10^{-5}	0.0012
2009	1	0.17648	-7.8×10^{-5}	-0.0015	0.000101	0.0015

3. 结果分析

首先采用指数平滑法对该航材进行需求预测,平滑系数 $\alpha = 0.8$,用一次指数平滑法进行预测。由于时间序列的数据较少,在 20 个以下时,初始值对以后

的预测影响很大。因此，选择以最初 2 期实际值的平均值作为初始值。

由图 7-10 可知，采用一次指数平滑法进行预测，预测曲线与实际数据的拟合性能不是很好，经计算得到 $MAE = 1.468$ ， $MSE = 2.654$ ，2009 年的预测需求量为 10.569。

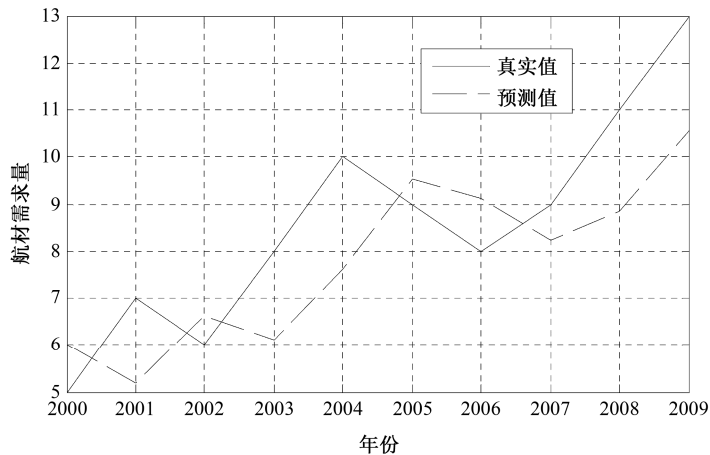


图 7-10 航材需求量预测——一次指数平滑法

运用支持向量机回归原理进行该航材的需求量预测。选择 RBF 核函数作为核函数，因为 RBF 核函数只有一个 g ，再加上一个惩罚系数 C ，所以只需对这两个参数进行优化。选择网格搜索法对惩罚系数 C 和核参数 g 进行优化，参数寻优的过程如图 7-11 所示。当 $C = 1024$ ， $g = 0.0315$ 时， MSE 取值达到最小值，此时的 (C, g) 取值为最优值。

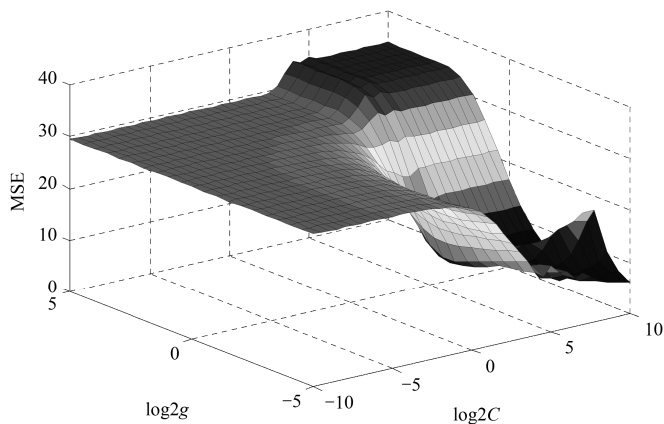


图 7-11 网格搜索法对支持向量机参数寻优

由图 7-12 可知，预测曲线与实际数据的拟合性能较好，经计算得到 $MAE = 0.654$ ， $MSE = 0.7191$ ，2009 年的需求预测值为 12.316。

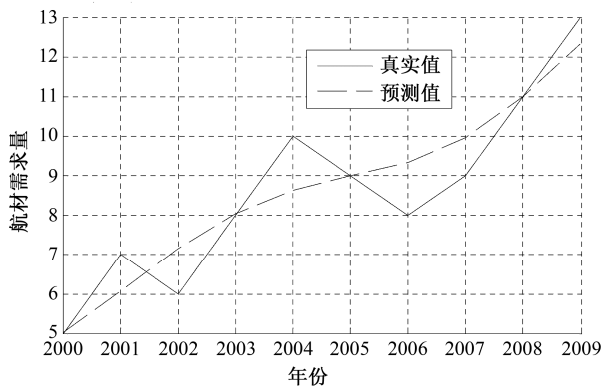


图 7-12 航材需求量预测——SVM—RBF 核函数

再次运用支持向量机回归原理对该航材进行需求量预测。选择 RBF 核函数作为核函数，参数优化的方法改为遗传算法，经优化参数 $C = 121.0345$ ， $g = 0.4983$ 。

由图 7-13 可知，预测曲线与实际数据的拟合性能很好，经计算得到 $MAE = 0.533$ ， $MSE = 0.5184$ ，2009 年的预测需求量为 12.457。

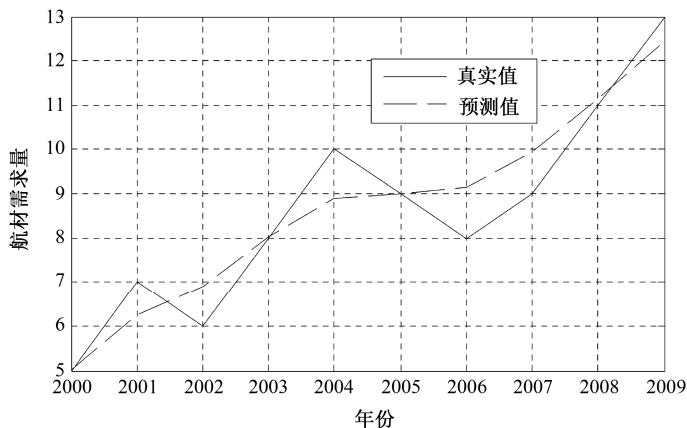


图 7-13 航材需求量预测——GA—SVM—RBF 核

由表 7-7 可知，对比三种方法的 MAE、MSE 值及 2009 年的需求预测值可知，采用遗传算法来优化参数的支持向量机进行预测所得到的结果和实际值最为

接近。由此可知，指数平滑法的预测效果较差，采用网格搜索法来优化参数的支持向量机的预测效果很好，采用遗传算法来优化参数的支持向量机预测性能最好。

表 7-7 预测结果对比

预测方法	参数优化方法	MAE	MSE	2009 年预测值	2009 年实际值
指数平滑法 ($\alpha = 0.8$)		1.468	2.456	10.569	13
支持向量机	网格搜索法	0.654	0.7191	12.316	13
	遗传算法	0.533	0.5184	12.457	13

7.3 基于复杂时间序列的组合相关向量机备件需求预测方法

针对武器装备备件需求预测面临的建模难度大、数据样本少、不确定性强等问题及典型机器学习方法存在的不足，本节研究了基于相关向量机的武器装备备件需求预测方法。

7.3.1 相空间重构

利用相关向量机特征参数序列进行预测，首先需要对备件消耗数据序列进行相空间重构。相空间重构的目的是找到相关向量机输入输出之间的对应关系。对于时间序列 $\{s_1, s_2, \dots, s_l\}$ ，假设预测目标值 s_n 与先前的 d 个值 $s_{n-1}, s_{n-2}, \dots, s_{n-d}$ 之间存在非线性映射关系：

$$s_n = f(s_{n-1}, s_{n-2}, \dots, s_{n-d}) \quad (n = d+1, d+2, \dots, l) \quad (7-32)$$

式中： d 为嵌入维数； $f(\bullet)$ 为非线性映射函数。

可以得到相关向量机的训练样本集：

$$X = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} s_1 & s_2 & \cdots & s_d \\ s_2 & s_3 & \cdots & s_{d+1} \\ \vdots & \vdots & & \vdots \\ s_N & s_{N+1} & \cdots & s_{d+N-1} \end{bmatrix}, \quad \mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_N \end{bmatrix} = \begin{bmatrix} s_{d+1} \\ s_{d+2} \\ \vdots \\ s_{d+N} \end{bmatrix}$$

式中： X 为输入样本； \mathbf{t} 为输出样本； $N = l - d$ 为训练样本数目。

利用该训练样本集训练相关向量机，则可以使训练后的相关向量机逼近

$f(\bullet)$ ，进而实现对 $\{s_1, s_2, \dots, s_l\}$ 后续数据的预测。

7.3.2 小波变换基本理论

小波分析的概念是 1984 年由法国地球物理学家 Morlet 在分析地震信号时提出来的。小波分析主要研究函数的表示，即将函数分解为“基本函数”之和，而“基本函数”是由一个小波函数经伸缩和平移得到的，这个小波函数具有很好的局部性和光滑性，使得人们通过分解系数表述函数时，可以分析函数的局部性质和整体性质。小波分析出现之前，人们用 Fourier 基、Haar 基分解函数。Fourier 基具有很好的光滑性，但局部性很差；而 Haar 基的局部性虽然好，但光滑性很差。小波基兼具两者的优点。

1. 连续小波变换

定义 7.1 设函数 $\psi \in L^2(\mathbf{R}) \cap L^1(\mathbf{R})$ ，并且 $\hat{\psi}(0) = 0$ ，由 ψ 经伸缩和平移得到一族函数：

$$\psi_{a,b}(x) = |a|^{-1/2} \psi\left(\frac{x-b}{a}\right) \quad (a, b \in \mathbf{R}, a \neq 0) \quad (7-33)$$

式中： $\{\psi_{a,b}\}$ 为分析小波或连续小波； ψ 为基本小波或母小波； a 为伸缩因子； b 为伸缩因子； $\hat{\psi}$ 代表 ψ 的傅里叶变换。

定义 7.2 设 ψ 为基本小波， $\{\psi_{a,b}\}$ 是由式 (7-33) 定义的连续小波，对于信号 $f \in L^2(\mathbf{R})$ ，其连续小波变换定义为

$$(W_\psi f)(a, b) = \langle f, \psi_{a,b} \rangle = |a|^{-1/2} \int_{-\infty}^{+\infty} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (7-34)$$

式中： $\overline{\psi(t)}$ 为 $\psi(t)$ 的复共轭。

定义 7.3 设 $\psi \in L^2(\mathbf{R}) \cap L^1(\mathbf{R})$ ，且满足条件

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\omega)|}{|\omega|} d\omega < \infty \quad (7-35)$$

则 ψ 称为允许小波，式 (7-35) 称为允许条件。

下面的定理 7.1 证明，如果基本小波 $\psi(t)$ 是允许小波，则可以从连续小波变换中恢复出原始信号。

定理 7.1 设 ψ 为允许小波，对任意的 $f, g \in L^2(\mathbf{R})$ ，有

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[(W_\psi f)(a, b) \overline{(W_\psi g)(a, b)} \right] \frac{da}{a^2} db = C_\psi \langle f, g \rangle \quad (7-36)$$

此外, 如果 f 在 $x \in \mathbf{R}$ 处连续, 则有

$$f(x) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (W_\psi f)(a, b) \left\{ |a|^{-1/2} \psi \left(\frac{x-b}{a} \right) \right\} \frac{da}{a^2} db \quad (7-37)$$

在信号分析中, 一般只考虑正频率, 即只考虑 a 取正值的情况。因此, 在重构信号 f 的过程中, 只能使用值 $(W_\psi f)(a, b)$ ($a > 0, b \in \mathbf{R}$)。这时对允许条件作进一步的加强, 即

$$\int_0^{+\infty} \frac{|\hat{\psi}(\omega)|}{|\omega|} d\omega = \int_0^{+\infty} \frac{|\hat{\psi}(-\omega)|^2}{|\omega|} d\omega = \frac{1}{2} C_\psi < \infty \quad (7-38)$$

定理 7.2 设 ψ 为满足式 (7-38) 的允许小波, 对任意的 $f, g \in L^2(\mathbf{R})$, 有

$$\int_0^{+\infty} \left[\int_{-\infty}^{+\infty} (W_\psi f)(a, b) \overline{(W_\psi g)(a, b)} db \right] \frac{da}{a^2} = \frac{1}{2} C_\psi \langle f, g \rangle \quad (7-39)$$

此外, 如果 f 在 $x \in \mathbf{R}$ 处连续, 则有

$$f(x) = \frac{2}{C_\psi} \int_0^{+\infty} \left[\int_{-\infty}^{+\infty} (W_\psi f)(a, b) \psi_{a,b}(x) db \right] \frac{da}{a^2} \quad (7-40)$$

2. 离散小波变换

在信号处理中, 一般采用离散小波变换。将式 (7-33) 中的参数 a 和 b 都取离散值, 固定伸缩步长 $a_0 > 1$, 位移步长 $b_0 \neq 0$, 取 $a = a_0^{-m}$, $b = nb_0 a_0^{-m}$, 从而将连续小波变为离散小波, 即

$$\psi_{m,n}(x) = a_0^{m/2} \psi(a_0^m x - nb_0) \quad (m, n \in \mathbf{Z}) \quad (7-41)$$

定义 7.4 设 ψ 为基本小波, 称式 (7-41) 定义的函数族 $\{\psi_{m,n}\}_{m,n \in \mathbf{Z}}$ 为离散小波, 对于信号 $f \in L^2(\mathbf{R})$, 其离散小波变换定义为

$$(Df)_{m,n} = \langle f, \psi_{m,n} \rangle = a_0^{m/2} \int_{-\infty}^{+\infty} f(t) \overline{\psi(a_0^m t - nb_0)} dt \quad (7-42)$$

如果 ψ 为允许小波, 并且有足够的衰减, 则 D 是把 $L^2(\mathbf{R})$ 映射到 $l^2(\mathbf{Z}^2)$ 。通常, D 不存在逆变换。如果 $\{\psi_{m,n}, m, n \in \mathbf{Z}\}$ 是 $L^2(\mathbf{R})$ 中的一个框架, 即存在 $0 < A \leq B < +\infty$, 有

$$A \|f\|_2^2 \leq \sum_{m,n \in \mathbf{Z}} |\langle f, \psi_{m,n} \rangle|^2 \leq B \|f\|_2^2 \quad (\forall f \in L^2(\mathbf{R})) \quad (7-43)$$

则可以建立从小波系数 $\langle f, \psi_{m,n} \rangle$ 重建 f 的数字方法。

3. 多分辨分析与 Mallat 算法

小波分析能够提供 $L^2(\mathbf{R})$ 中具有良好局部化性质的正交基, 把 $L^2(\mathbf{R})$ 中的函数与 $l^2(\mathbf{Z})$ 的数列等同起来, 从而把分析问题转化为代数问题来解决。多分辨分析的思想是: 首先在能量有限函数空间 $L^2(\mathbf{R})$ 的某个子空间建立基底; 然后利用简单的伸缩和平移变换, 把子空间的基底扩充到 $L^2(\mathbf{R})$ 中。

定义 7.5 设 $\{V_j\}_{j \in \mathbf{Z}}$ 是 $L^2(\mathbf{R})$ 空间中的一个闭子空间, $\{V_j\}_{j \in \mathbf{Z}}$ 称为多分辨分析, 如果 $\{V_j\}_{j \in \mathbf{Z}}$ 满足如下四个条件:

- (1) $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$ 。
- (2) $\bigcap_{j \in \mathbf{Z}} V_j = \{0\}$, $\overline{\bigcup_{j \in \mathbf{Z}} V_j} = L^2(\mathbf{R})$ 。
- (3) $\forall j \in \mathbf{Z}$, 若 $f(x) \in V_j$, 则 $f(2x) \in V_{j+1}$ 。
- (4) 存在 $\varphi \in V_0$, 使得 $\{\varphi(x-n)\}_{n \in \mathbf{Z}}$ 是 V_0 的 Riesz 基, 即
 - ① $V_0 = \overline{\text{span}\{\varphi(x-n)\}_{n \in \mathbf{Z}}}$;
 - ② 存在 $0 < A \leq B < +\infty$, 使得对任意 $\{c_n\}_{n \in \mathbf{Z}} \in l^2(\mathbf{Z})$, 有

$$A \sum_{n \in \mathbf{Z}} |c_n|^2 \leq \left\| \sum_{n \in \mathbf{Z}} c_n \varphi(x-n) \right\|_2^2 \leq B \sum_{n \in \mathbf{Z}} |c_n|^2 \quad (7-44)$$

式中: φ 为尺度函数。称满足式 (7-44) 的函数族 $\{\varphi(x-n)\}_{n \in \mathbf{Z}}$ 为 l^2 稳定的。多分辨分析又称为多尺度分析。

条件 (1) 称为一致单调性, 表明子空间 $\{V_j\}_{j \in \mathbf{Z}}$ 是嵌套的。条件 (2) 称为渐近完全性, 表明通过增大 j , $L^2(\mathbf{R})$ 中的每一个函数 f 能够用它在 V_j 中的投影 $P_j f$ 非常接近所希望的逼近; 反之, 通过减小 j , 投影 $P_j f$ 能够具有任意小的能量。条件 (3) 称为伸缩规则性, 表明空间列 $\{V_j\}_{j \in \mathbf{Z}}$ 中任一空间 V_l 的基可由其中另一个空间 V_j 的基经过伸缩与平移交换得到。条件 (4) 称为 Riesz 基的存在性, 结合条件 (3) 可知, 由 $\{\varphi(x-n)\}_{n \in \mathbf{Z}}$ 是 V_0 的 Riesz 基, 可得 $\{2^{j/2} \varphi(2^j x - k)\}_{k \in \mathbf{Z}}$ 构成 V_j 的 Riesz 基。因此, φ 又称为多分辨分析 $\{V_j\}_{j \in \mathbf{Z}}$ 的生成元。

定理 7.3 设 $\{V_j\}_{j \in \mathbf{Z}}$ 是 $L^2(\mathbf{R})$ 中的一个多分辨率分析, 则存在函数 $\varphi \in V_0$, 使得 $\{\varphi(x-n)\}_{n \in \mathbf{Z}}$ 是 V_0 的规范正交基, 从而 $\{2^{j/2} \varphi(2^j x - k)\}_{k \in \mathbf{Z}}$ 构成 V_j 的规范正交基。

多分辨分析能够提供低通滤波器和高通滤波器，可用于信号的分解与重构。

因为 $\frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right) \in V_{-1} \subset V_0$ ，故存在 $\{h_k\}_{k \in \mathbf{Z}} \in l^2(\mathbf{Z})$ ，有双尺度方程

$$\frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right) = \sum_k h_k \varphi(x-k) \quad (7-45)$$

由 $\{\varphi(x-n)\}_{n \in \mathbf{Z}}$ 是 V_0 的规范正交基，可得

$$h_k = \left\langle \frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right), \varphi(x-k) \right\rangle = \int_{\mathbf{R}} \frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right) \overline{\varphi(x-k)} dx \quad (7-46)$$

$\{h_k\}_{k \in \mathbf{Z}}$ 就是所需要的低通滤波器的冲击响应。

下面定理构造了高通滤波器的冲击响应、正交小波函数与正交小波基。

定理 7.4 设 $\{V_j\}_{j \in \mathbf{Z}}$ 是 $L^2(\mathbf{R})$ 中的一个多分辨率分析， φ 为尺度函数，满足：

(1) $\{\varphi(x-n)\}_{n \in \mathbf{Z}}$ 是 V_0 的规范正交基。

(2) 存在 $\{h_k\}_{k \in \mathbf{Z}} \in l^2(\mathbf{Z})$ ，有 $\frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right) = \sum_k h_k \varphi(x-k)$ 成立。定义序列 $\{g_k\}_{k \in \mathbf{Z}} \in l^2(\mathbf{Z})$ ，函数 $\psi(x)$ 与空间 W_j 如下式：

$$g_k = (-1)^k \overline{h_{-k}} \quad (7-47)$$

$$\frac{1}{\sqrt{2}}\psi\left(\frac{x}{2}\right) = \sum_k g_k \varphi(x-k) \quad (7-48)$$

$$W_j = \overline{\text{span}\{\psi_{j,n}(x) = 2^{j/2}\psi(2^j x - n), n \in \mathbf{Z}\}} \quad (7-49)$$

则有：① $W_j \perp V_j$ ， $V_{j+1} = W_j \oplus V_j$ ，从而， $L^2(\mathbf{R}) = \bigoplus_{j \in \mathbf{Z}} \overline{W_j}$ ， $W_i \perp W_j$ ， $i \neq j$ ；②

$\{\psi_{j,n}\}_{n \in \mathbf{Z}}$ 是 W_j 的规范正交基，从而 $\{\psi_{j,n}\}_{j,n \in \mathbf{Z}}$ 是 $L^2(\mathbf{R})$ 的规范正交基。

$\psi(x)$ 称为小波函数， $\{\psi_{j,n}\}_{j,n \in \mathbf{Z}}$ 称为小波基。此时， $\psi(x)$ 为正交小波函数，

$\{\psi_{j,n}\}_{j,n \in \mathbf{Z}}$ 为正交小波基，因为 $\{\psi_{j,n}\}_{j,n \in \mathbf{Z}}$ 构成 $L^2(\mathbf{R})$ 的正交基。 \perp 表示正交运算，

\oplus 表示正交和运算。

由 $V_{j+1} = W_j \oplus V_j$ 可得，对任意整数 N 与 $M(>0)$ 有 $V_N = W_{N-1} \oplus V_{N-1} = W_{N-1} \oplus W_{N-2} \oplus \cdots \oplus W_{N-M} \oplus V_{N-M}$ 。从而， $\forall f_N \in V_N$ ，有唯一的分解，即

$$f_N = e_{N-1} + f_{N-1} = e_{N-1} + e_{N-2} + \cdots + e_{N-M} + f_{N-M}$$

式中： $f_j \in V_j$ ， $e_j \in W_j$ 。

$\forall f_j \in V_j, e_j \in W_j$, 它们有唯一的级数表示, 即

$$\begin{cases} f_j(x) = \sum_k c_k^j \varphi_{j,k}(x) \\ C^j = \{c_k^j\}_{k \in \mathbf{Z}} \in l^2(\mathbf{Z}) \end{cases} \quad (7-50)$$

$$\begin{cases} e_j(x) = \sum_k d_k^j \psi_{j,k}(x) \\ D^j = \{d_k^j\}_{k \in \mathbf{Z}} \in l^2(\mathbf{Z}) \end{cases} \quad (7-51)$$

由 $f_{j+1} = e_j + f_j$ 可得

$$f_{j+1}(x) = \sum_k c_k^{j+1} \psi_{j+1,k}(x) = \sum_k c_k^j \psi_{j,k}(x) + \sum_k d_k^j \psi_{j,k}(x) \quad (7-52)$$

从而有

$$\begin{aligned} c_k^j &= \sum_{l \in \mathbf{Z}} c_l^{j+1} \langle \varphi_{j+1,l}(x), \varphi_{j,k}(x) \rangle \\ d_k^j &= \sum_{l \in \mathbf{Z}} c_l^{j+1} \langle \varphi_{j+1,l}(x), \psi_{j,k}(x) \rangle \\ c_k^{j+1} &= \sum_{l \in \mathbf{Z}} c_l^j \langle \varphi_{j,l}(x), \varphi_{j+1,k}(x) \rangle + \sum_{l \in \mathbf{Z}} d_l^j \langle \psi_{j,l}(x), \varphi_{j+1,k}(x) \rangle \end{aligned}$$

经计算可得 Mallat 分解算法和重构算法:

$$\begin{cases} c_k^j = \sum_{l \in \mathbf{Z}} c_l^{j+1} \bar{h}_{l-2k} \\ d_k^j = \sum_{l \in \mathbf{Z}} c_l^{j+1} \bar{g}_{l-2k} \end{cases} \quad (7-53)$$

$$c_k^{j+1} = \sum_{l \in \mathbf{Z}} c_l^j h_{k-2l} + \sum_{l \in \mathbf{Z}} d_l^j g_{k-2l} \quad (7-54)$$

式中: $\{h_k\}_{k \in \mathbf{Z}} \in l^2(\mathbf{Z})$ 由式 (7-46) 给出, 可看成低通滤波器系数; $\{g_k\}_{k \in \mathbf{Z}} \in l^2(\mathbf{Z})$ 由式 (7-47) 给出, 可看成高通滤波器系数; c_k^j 是 c_l^{j+1} 的近似信号; d_k^j 是 c_l^{j+1} 的细节信号。

根据 Mallat 算法, 一维信号的小波分解与重构过程可由两组滤波器的级联滤波而产生。设 $\{c_k^m\}$ 为输入序列, $\{c_k^{m-i}\}$ 为经 i 次低通滤波器而得到的输出, $\{d_k^{m-i}\}$ 为第 i 次高通滤波器得到的输出, 一维信号的小波分解与重构过程分别如图 7-14 和图 7-15 所示。这种分解方式称为塔式分解, 其中, 每次都需对低频分量 (近似信号) 进行分解。

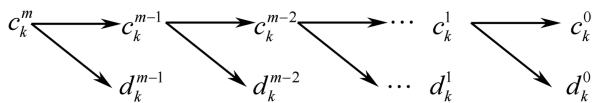


图 7-14 小波分解过程

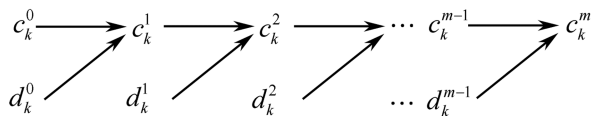


图 7-15 小波重构过程

7.3.3 小波函数的选取

在利用小波变换方法对装备性能特征参数序列处理的过程中,小波函数的选择十分重要。对于同一序列使用不同的小波函数对其进行分解,将得到不同的效果。因此,选择一个合适的小波函数是利用组合相关向量机进行故障预测的前提和基础。目前,关于小波函数如何选取的问题还没有完善的理论作为指导。根据组合相关向量机的建模过程,本书在选择小波函数时,主要考虑小波函数的以下特性:

(1) 正交性。为根据小波变换后各分量的预测均值和方差得到最终的预测均值和方差,需要保证波变换后的各分量是相互独立的。由于各分量在建立相关向量机回归预测模型时是假设服从高斯分布的,因此,只需要满足各分量是互不相关的。对于小波变换不同尺度分解结果的正交性,可以证明有如下的定理:

定理 7.5 在小波变换中,如果 $\psi(x)$ 为正交小波函数,当尺度相异时,则此二尺度的分解结果互不相关、彼此正交;而同一尺度的分解结果的自相关性与原始信号的自相关性相同。

因此,在选择小波函数时应选择具有正交性的小波函数。

(2) 正则性。正则性一般用来描述函数的光滑程度,正则性越高,函数的光滑性越好。小波函数的正则性主要影响小波系数重构的稳定性,通常对小波要求一定的正则性是为了获得更好的重构信号。因此,正则性是选择小波函数的另一个重要依据。如果小波函数的正则性系数大,则此小波函数能够较好地反映信号的整体变化趋势。

(3) 紧支撑性。紧支撑是小波的重要性质,良好的紧支撑性能够凸显信号的

时域特性。紧支撑越短，小波的局部化特性越好，小波变换的计算复杂度越低，从而更便于快速实现。

目前，常用的小波主要有 Haar 小波、Meyer 小波、Symlets 小波和 Daubechies 小波等。Daubechies 小波具有正交、时频紧支撑、高正则性等特点，在处理非平稳时间序列时具有很好的特性，因此采用 Daubechies 小波。

7.3.4 基于组合相关向量机的预测原理

7.3.4.1 相关向量机的预测原理

相关向量机的概念是 2000 年由美国学者 M. E. Tipping 提出的。回归用于预测输入变量（自变量）和输出变量（因变量）之间的关系，特别是当输入变量的值发生变化时，输出变量的值随之发生的变化。回归模型正是表示输入变量到输出变量之间映射的函数。回归问题的学习等价于函数拟合：选择一条函数曲线使其能够很好地拟合已知数据且很好地预测未知数据。

回归问题分为学习和预测两个过程。首先给定训练数据集 $\{\mathbf{x}_i, t_i\}_{i=1}^N$ ($\mathbf{x}_i \in \mathbf{R}^d$ 为样本输入， $t_i \in \mathbf{R}$ 为对应的输出)。学习系统基于训练数据构建一个模型，即函数 $\mathbf{Y} = f(\mathbf{X})$ ；对新的输入 \mathbf{x}_{N+1} ，预测系统根据学习的模型确定相应的输出 t_{N+1} 。

给定训练样本集 $\{\mathbf{x}_n, t_n\}_{n=1}^N$ ($\mathbf{x}_n \in \mathbf{R}^d$ ， $t_n \in \mathbf{R}$)，相关向量机故障预测模型输出可以表示为

$$y(\mathbf{x}, \mathbf{w}) = \sum_{i=1}^N w_i K(\mathbf{x}, \mathbf{x}_i) + w_0 \quad (7-55)$$

式中： $K(\mathbf{x}, \mathbf{x}_i)$ 为核函数； w_i ($i = 0, 1, 2, \dots, N$) 为模型的权值； N 为训练样本数。

假定目标值是独立的，且来自带噪声的模型：

$$t_n = y(\mathbf{x}_n, \mathbf{w}) + \varepsilon_n \quad (7-56)$$

式中： ε_n 为噪声，服从均值为 0、方差为 σ^2 的高斯分布。

相应的训练样本集的似然函数为

$$p(\mathbf{t} | \mathbf{w}, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2} \|\mathbf{t} - \Phi\mathbf{w}\|^2\right\} \quad (7-57)$$

式中：训练样本集的目标值向量 $\mathbf{t} = [t_1, \dots, t_N]^T$ ；模型权值向量 $\mathbf{w} = [w_0, w_1, \dots, w_N]^T$ ； Φ 是 $N \times (N+1)$ 的构造矩阵且 $\Phi = [\phi(\mathbf{x}_1), \phi(\mathbf{x}_2), \dots, \phi(\mathbf{x}_N)]^T$ ， $\phi(\mathbf{x}_n) = [1, K(\mathbf{x}_n, \mathbf{x}_1), \dots, K(\mathbf{x}_n, \mathbf{x}_N)]^T$ 。

通过最大似然法求出的最优 \mathbf{w} 能够使经验风险最小化，但会导致过学习。为了避免这种情况，采用稀疏贝叶斯方法对权值 \mathbf{w} 赋予先验的条件概率分布：

$$p(\mathbf{w} | \boldsymbol{\alpha}) = \prod_{i=0}^N N(w_i | 0, \alpha_i^{-1}) \quad (7-58)$$

式中： $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \dots, \alpha_N]^T$ 是 $N+1$ 维的超参数向量，每一个独立的 α_i 都只与其对应的权值 w_i 相关。

引入超参数 $\beta = \sigma^{-2}$ ，定义 $\boldsymbol{\alpha}$ 和 β 的超先验 $p(\boldsymbol{\alpha})$ 、 $p(\beta)$ 为伽马分布：

$$p(\boldsymbol{\alpha}) = \text{Gamma}(\boldsymbol{\alpha} | a, b) = \prod_{i=0}^N \frac{b^a}{\Gamma(a)} \alpha_i^{a-1} e^{-b\alpha_i}$$

$$p(\beta) = \text{Gamma}(\beta | c, d) = \frac{d^c}{\Gamma(c)} \beta^{c-1} e^{-d\beta}$$

式中： $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$ ； a 、 b 、 c 、 d 可以设定为很小的值，本节设 $a=b=c=d=0$ 。

根据贝叶斯公式，对所有未知参数有如下后验公式：

$$p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 | \mathbf{t}) = \frac{p(\mathbf{t} | \mathbf{w}, \boldsymbol{\alpha}, \sigma^2) p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2)}{p(\mathbf{t})} \quad (7-59)$$

然后，给定新观测点 \mathbf{x}_* ，相应目标预测值 t_* 分布为

$$p(t_* | \mathbf{t}) = \int p(t_* | \mathbf{w}, \boldsymbol{\alpha}, \sigma^2) p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 | \mathbf{t}) d\mathbf{w} d\boldsymbol{\alpha} d\sigma^2 \quad (7-60)$$

式 (7-60) 不能直接计算，考虑

$$p(\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 | \mathbf{t}) = p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{t}) \quad (7-61)$$

则有

$$p(\mathbf{w} | \mathbf{t}, \boldsymbol{\alpha}, \sigma^2) = \frac{p(\mathbf{t} | \mathbf{w}, \sigma^2) p(\mathbf{w} | \boldsymbol{\alpha})}{p(\mathbf{t} | \boldsymbol{\alpha}, \sigma^2)} \quad (7-62)$$

$$= (2\pi)^{-(N+1)/2} |\boldsymbol{\Sigma}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{w} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{w} - \boldsymbol{\mu}) \right\}$$

式中： $\boldsymbol{\Sigma}$ 为后验协方差矩阵 $\boldsymbol{\mu}$ 为后验均值向量。它们分别为

$$\boldsymbol{\Sigma} = (\sigma^{-2} \boldsymbol{\Phi}^T \boldsymbol{\Phi} + \mathbf{A})^{-1} \quad (7-63)$$

$$\boldsymbol{\mu} = \sigma^{-2} \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \mathbf{t} \quad (7-64)$$

式中： $\mathbf{A} = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_N)$ 。

对式 (7-61) 中超参数后验 $p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{t})$ ，用 δ 函数的峰值逼近。为了预测，不要求 $p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{t}) \approx \delta(\boldsymbol{\alpha}_{\text{MP}}, \sigma_{\text{MP}}^2)$ ，仅要求

$$\int p(t_* | \mathbf{w}, \boldsymbol{\alpha}, \sigma^2) \delta(\boldsymbol{\alpha}_{\text{MP}}, \sigma_{\text{MP}}^2) d\boldsymbol{\alpha} d\sigma^2 \approx \int p(t_* | \boldsymbol{\alpha}, \sigma^2) p(\boldsymbol{\alpha}, \sigma^2 | \mathbf{t}) d\boldsymbol{\alpha} d\sigma^2 \quad (7-65)$$

是一个好的估计。于是，RVM 学习问题就转化为寻求超参数后验分布 $p(\alpha, \sigma^2 | t) \propto p(t | \alpha, \sigma^2) p(\alpha) p(\sigma^2)$ 关于 α 和 σ^2 的最大值问题。在一致超先验情况下，仅需求 $p(t | \alpha, \sigma^2)$ 最大值：

$$\begin{aligned} p(t | \alpha, \sigma^2) &= \int p(t | w, \sigma^2) p(w | \alpha) dw \\ &= (2\pi)^{-N/2} |\sigma^2 \mathbf{I} + \Phi A^{-1} \Phi^T|^{-1/2} \cdot \exp \left\{ -\frac{1}{2} t^T (\sigma^2 \mathbf{I} + \Phi A^{-1} \Phi^T)^{-1} t \right\} \end{aligned} \quad (7-66)$$

其等价于最大化其对数函数 $L(\alpha)$ ：

$$\begin{aligned} L(\alpha) &= \log p(t | \alpha, \sigma^2) = \log \int_{-\infty}^{+\infty} p(t | w, \sigma^2) p(w | \alpha) dw \\ &= -\frac{1}{2} \left[N \log 2\pi + \log |\mathbf{C}| + t^T \mathbf{C}^{-1} t \right] \end{aligned} \quad (7-67)$$

式中： $\mathbf{C} = \sigma^2 \mathbf{I} + \Phi A^{-1} \Phi^T$ 。

通过最大化边缘似然函数 $p(t | \alpha, \sigma^2)$ 估计 α 和 σ^2 的方法称为第 II 类型边缘似然最大化方法。

获得参数 α_i^{new} 后，重新估计权值的后验均值和方差。在迭代估计过程中，大部分 α_i 值会越来越接近于无穷大，即对应的 w_i 为 0，其相应的基函数可以删除，从而达到稀疏性。其余的 α_i 会趋近有限值，与之对应的 x_i 即称为相关向量。

假设边缘似然最大化估计值为 α_{MP} 和 σ_{MP}^2 ，若给定输入值 x_* ，则相应输出的概率分布为

$$p(t_* | t, \alpha_{\text{MP}}, \sigma_{\text{MP}}^2) = \int p(t_* | w, \sigma_{\text{MP}}^2) p(w | t, \alpha_{\text{MP}}, \sigma_{\text{MP}}^2) dw \quad (7-68)$$

式 (7-68) 服从正态分布，即

$$p(t_* | t, \alpha_{\text{MP}}, \sigma_{\text{MP}}^2) = N(t_* | y_*, \sigma_*^2)$$

其中，预测均值和方差分别为

$$y_* = \mu^T \phi(x_*) \quad (7-69)$$

$$\sigma_*^2 = \sigma_{\text{MP}}^2 + \phi(x_*)^T \Sigma \phi(x_*) \quad (7-70)$$

由式 (7-70) 可以看出，相关向量机预测的方差由两部分组成：一是数据噪声；二是模型权值估计的不确定性引起的预测方差。

7.3.4.2 组合相关向量机建模的基本过程

组合相关向量机建模的基本过程：首先，通过正交小波变换将装备性能特征参数序列分解为具有不同特征尺度的随机分量和趋势分量，降低建模数据的复杂度；其次，分别建立各随机分量和趋势分量的相关向量机回归预测模型，并采用改进的量子粒子群优化算法优化选择各分量预测模型的核参数和嵌入维数；最

后, 叠加各相关向量机回归预测模型的预测均值和方差作为最终的预测均值和方差, 进而得到预测值的概率分布。组合相关向量机的预测框架如图 7-16 所示。

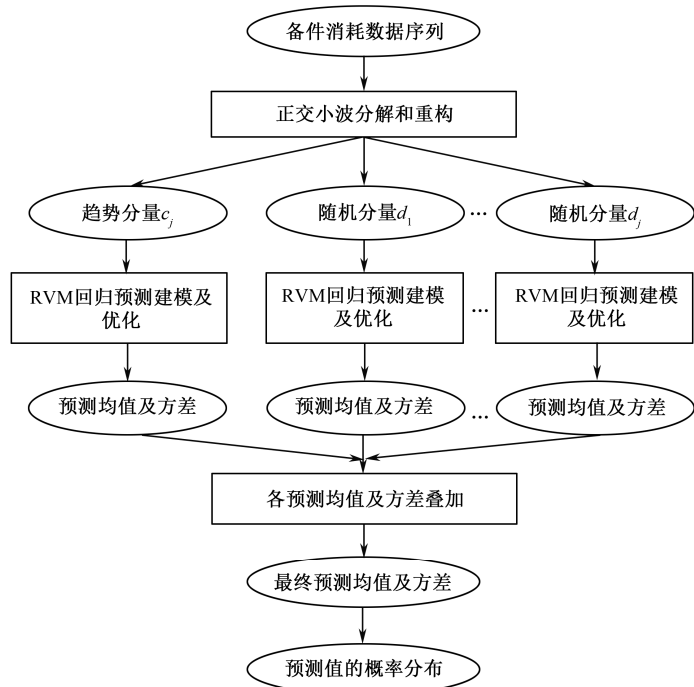


图 7-16 组合相关向量机的预测框架

7.4.4.3 基于量子粒子群算法的相关向量机模型优化

1. 量子粒子群算法

传统的 PSO 算法将群体中的每个个体（可行解）看作搜索空间中的一个没有体积的粒子，在搜索空间中以一定的速度飞行，其速度根据目前所找到的最优位置和整个粒子群所找到的最优位置进行动态调整。量子粒子群优化（Quantum-behaved Particle Swarm Optimization, QPSO）算法结合量子物理的思想对 PSO 算法中粒子速度和位置的更新方法进行了改进。QPSO 算法中粒子只有位置向量，没有速度向量，在更新粒子位置时考虑了各个粒子的当前局部最优位置和全局最优位置信息。

1) 粒子的基本进化方程

在 N 维的目标搜索空间中，种群中粒子数目为 M ， t 时刻第 i ($i=1,2,\dots,M$)

个粒子位置为 $\mathbf{X}_i(t)=[X_{i1}(t), X_{i2}(t), \dots, X_{iN}(t)]$ ，每个粒子的个体最优位置 $\mathbf{P}_i(t)=[P_{i1}(t), P_{i2}(t), \dots, P_{iN}(t)]$ ，群体的全局最优位置 $\mathbf{P}_g(t)=[P_{g1}(t), P_{g2}(t), \dots, P_{gN}(t)]$ ，则第 i 个粒子第 j 维坐标的基本进化方程为

$$X_{i,j}(t+1) = P_{i,j}(t) \pm \frac{L_{i,j}(t)}{2} \ln(1/u_{i,j}) \quad (7-71)$$

式中： $u_{i,j}$ 为在 (0,1) 上均匀分布的随机数； $L_{i,j}$ 是以粒子位置坐标 $P_{i,j}$ 为中心建立的 δ 势阱的特征长度。

2) 搜索迭代策略

搜索迭代策略解决的关键问题是如何控制 $L_{i,j}$ ，使其能够收敛至 0，从而使粒子收敛到个体最优位置为 \mathbf{P}_i 。通常有两种方法：

(1) $L_{i,j}$ 由下式确定：

$$L_{i,j}(t) = 2\rho |P_{i,j}(t) - X_{i,j}(t)| \quad (7-72)$$

则粒子的进化方程变为

$$X_{i,j}(t+1) = P_{i,j}(t) \pm \rho |P_{i,j}(t) - X_{i,j}(t)| \ln[1/u_{i,j}(t)] \quad (7-73)$$

(2) 在算法中引入粒子的平均最优位置，记为 $\mathbf{C}(t)$ 。 $\mathbf{C}(t)$ 定义为所有粒子个体最优位置的平均，即

$$\begin{aligned} \mathbf{C}(t) &= [C_1(t), C_2(t), \dots, C_N(t)] = \frac{1}{M} \sum_{i=1}^M \mathbf{P}_i(t) \\ &= \left[\frac{1}{M} \sum_{i=1}^M P_{i,1}(t), \frac{1}{M} \sum_{i=1}^M P_{i,2}(t), \dots, \frac{1}{M} \sum_{i=1}^M P_{i,N}(t) \right] \end{aligned} \quad (7-74)$$

$L_{i,j}$ 由下式确定：

$$L_{i,j}(t) = 2\rho |C_j(t) - X_{i,j}(t)| \quad (7-75)$$

则粒子的进化方程变为

$$X_{i,j}(t+1) = P_{i,j}(t) \pm \rho |C_j(t) - X_{i,j}(t)| \ln[1/u_{i,j}(t)] \quad (7-76)$$

式 (7-72)、式 (7-73)、式 (7-75) 和式 (7-76) 中的 ρ 称为收缩-扩张因子。 ρ 为 QPSO 算法中除群体规模和迭代次数外唯一需要确定的参数，一般通过距离自适应调节的方法来控制 ρ 的取值：

$$\rho = 0.5 + 0.5 \times \frac{\max \text{ iter} - \text{iter}}{\max \text{ iter}} \quad (7-77)$$

式中： $\max \text{ iter}$ 为算法的最大迭代次数； iter 为当前迭代次数。

通过对测试函数和其他应用问题的结果比较表明：第二种方法的性能远优于第一种方法。

(3) QPSO 算法流程。

QPSO 算法的主要步骤如下：

① 设置当前迭代次数 $\text{iter} = 1$ ，设定算法的最大迭代次数 $\max \text{iter}$ ，种群中粒子数目 M ，在搜索区间中初始化粒子群中粒子的位置。

② 将每个粒子的个体最优位置 P_i 设置为当前位置，计算每个粒子的适应度函数，取适应度最好的粒子所对应的个体最优位置作为最初的全局最优位置 P_g 。

③ 根据式 (3-14)、式 (3-16)、式 (3-17) 更新粒子位置，其中

$$P_{i,j}(t) = v_j(t)P_{i,j}(t) + [1 - v_j(t)]P_{g,j}(t) \quad (7-78)$$

式中： v_j 为在 (0,1) 上均匀分布的随机数。

④ 计算当前粒子的适应度值，并与其经历的最优位置 P_i 进行比较。若优，则将其作为当前最优位置，更新 P_i ；否则，保留原值。

⑤ 将更新后每个粒子的 P_i 与全局最优位置 P_g 比较。若优，则将 P_g 设置为最好粒子的当前位置；否则，保留原值。

⑥ 判断是否满足终止条件。如果满足，则将 P_g 作为最终结果；如果不满足，则设定 $\text{iter} = \text{iter} + 1$ 转至步骤③。

2. 基于 QPSO 算法的相关向量机模型优化

学习方法的泛化能力是指由该方法学习到的模型对未知数据的预测能力。好的泛化能力估计方法是进行相关向量机模型优化的基础。目前，采用最多的办法是将通过交叉验证得到的测试误差来作为学习方法泛化能力估计。

1) 简单交叉验证

随机地训练样本集划分为两部分：一部分作为训练集，另一部分作为测试集。利用训练集训练模型，然后利用测试集测试模型，将得到的测试误差作为分类或回归的性能指标。该方法的不足在于：测试结果与训练样本集的分组有很大关系，结果缺乏说服力。

2) L 折交叉验证

将训练样本集划分为 L 个互不相交、大小相同的独立子集，将其中的一个子集作为测试集，利用剩余的 $L-1$ 个子集训练模型并进行测试。重复上述过程 L 次，每次选取不同的子集作为测试集，得到每个子集作为测试集时的测试误差，称所有测试误差的平均值为交叉验证误差，并将其作为分类或回归的性能指标。以 5 折交叉验证为例，其步骤如图 7-17 所示。

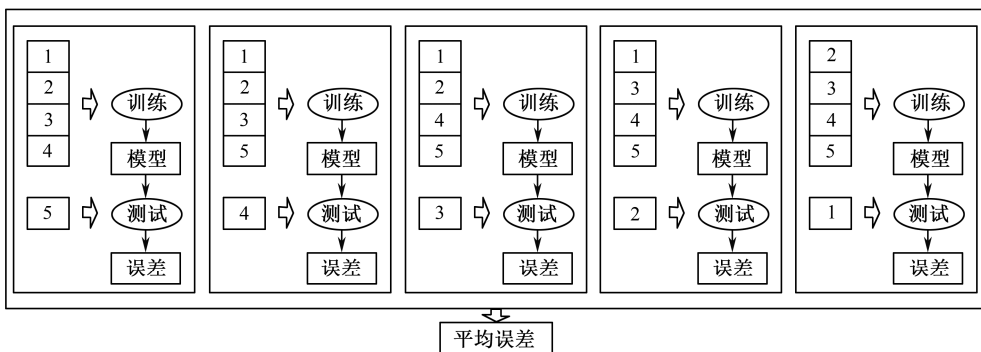


图 7-17 5 折交叉验证步骤

3. 留一交叉验证

假设训练样本集的样本数目为 N ，每次选取其中的一个样本作为测试集，利用剩余的 $N-1$ 个子集训练模型并进行测试。重复上述过程 N 次，每次选取不同的样本作为测试集，得到每个样本作为测试集时的测试误差，称所有测试误差的平均值为留一交叉验证误差，并将其作为分类或回归的性能指标。留一交叉验证是 L 折交叉验证的一个特例，当 $L=N$ 时， L 折交叉验证为留一交叉验证。

综上所述，本书以 L 折交叉验证得到的测试误差作为利用 QPSO 算法对相关向量机模型进行优化时的适应度函数。

4. 组合相关向量机模型的预测均值及方差

根据概率论多维随机变量及其分布的结论可知：一般，设 X_1 、 X_2 相互独立且 $X_1 \sim N(\mu_1, \sigma_1^2)$ ， $X_2 \sim N(\mu_2, \sigma_2^2)$ ，则 $Z = X_1 + X_2$ 仍然服从高斯分布，且有 $Z \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ 。将这个结论推广到 n 个独立高斯随机变量之和的情况，即若 $X_i \sim N(\mu_i, \sigma_i^2)$ ($i=1, 2, \dots, n$)，且它们相互独立，则它们的和 $Z = X_1 + X_2 + \dots + X_n$ 仍然服从高斯分布，且有 $Z \sim N(\mu_1 + \mu_2 + \dots + \mu_n, \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2)$ 。

小波变换各分量预测输出 t_i^* 服从均值为 y_i^* 、方差为 $(\sigma_i^*)^2$ 的高斯分布，即 $t_i^* \sim N(y_i^*, (\sigma_i^*)^2)$ ($i=1, 2, \dots, n$)，则有

$$t^* \sim N(y_1^* + y_2^* + \dots + y_n^*, (\sigma_1^*)^2 + (\sigma_2^*)^2 + \dots + (\sigma_n^*)^2) \quad (7-79)$$

记

$$y^* = y_1^* + y_2^* + \dots + y_n^* \quad (7-80)$$

$$\sigma^* = \sqrt{(\sigma_1^*)^2 + (\sigma_2^*)^2 + \cdots + (\sigma_n^*)^2} \quad (7-81)$$

则有

$$\frac{t^* - y^*}{\sigma^*} \sim N(0,1) \quad (7-82)$$

5. 组合相关向量机模型预测置信区间

下面给出 t_* 双侧置信区间的计算方法。假设给定置信度为 $1-\theta$ ，则有

$$P\left\{-z_{\theta/2} < \frac{t_* - y_*}{\sigma_*} < z_{\theta/2}\right\} = 1 - \theta \quad (7-83)$$

即

$$P\{y_* - \sigma_* z_{\theta/2} < t_* < y_* + \sigma_* z_{\theta/2}\} = 1 - \theta \quad (7-84)$$

于是得 t_* 的置信度为 $1-\theta$ 的置信区间为

$$[y_* - \sigma_* z_{\theta/2}, y_* + \sigma_* z_{\theta/2}] \quad (7-85)$$

式中：上分位数 $z_{\theta/2}$ 可通过查标准正态分布表得到。

同理，可以计算 t_* 的单侧置信区间：

$$(-\infty, y_* + \sigma_* z_{\theta}] \quad (7-86)$$

$$[y_* - \sigma_* z_{\theta}, +\infty) \quad (7-87)$$

式中： $y_* + \sigma_* z_{\theta}$ 为单侧置信上界； $y_* - \sigma_* z_{\theta}$ 为单侧置信下界。

7.3.5 示例分析

1. 试验数据

对于某一基地级后方仓库的某一种常用通用备件，对其月份历史消耗数据进行整理，见表 7-8 和图 7-18。以数据中的最后 5 个数据作为测试数据，其余数据作为建模数据，进一步预测试验。

表 7-8 备件消耗数据

序号	1	2	3	4	5	6	7	8	9	10
数值	39	29	29	33	32	334	39	31	29	30
序号	11	12	13	14	15	16	17	18	19	20
数值	32	32	38	38	33	35	32	36	36	36

续表

序号	21	22	23	24	25	26	27	28	29	30
数值	34	32	35	34	34	36	38	34	39	39
序号	31	32	33	34	35	36	37	38	39	40
数值	39	36	35	36	35	39	36	34	39	38
序号	41	42	43	44	45	46	47	48	49	50
数值	37	35	33	36	34	37	34	35	35	37
序号	51	52	53	54	55	56	57	58	59	60
数值	39	37	35	38	35	36	41	40	40	41
序号	61	62	63	64	65	66	67	68	69	70
数值	40	43	43	39	40	40	42	42	42	40
序号	71	72	73	74	75	76	77	78	79	80
数值	40	42	44	44	46	43	43	43	46	46
序号	81	82	83	84	85	86	87	88	89	90
数值	50	49	44	50	47	44	52	41	45	48
序号	91	92	93	94	95	96	97	98	99	100
数值	47	48	48	48	50	51	49	52	49	48
序号	101	102	103	104	105	106	107	108	109	110
数值	50	—	—	—	—	—	—	—	—	—

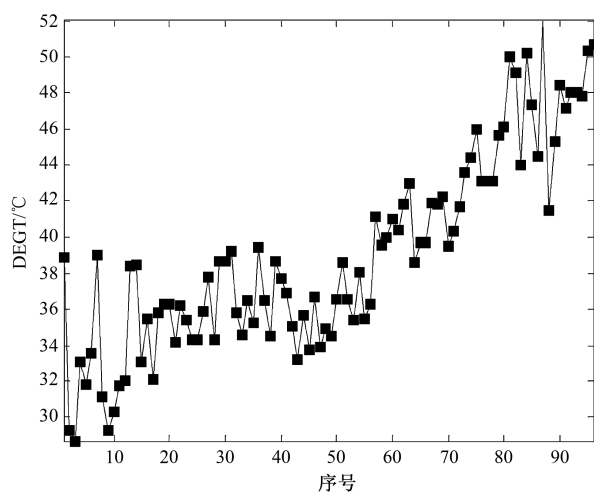


图 7-18 备件消耗数据

2. 小波变换结果

1 层小波分解与重构结果如图 7-19 所示。

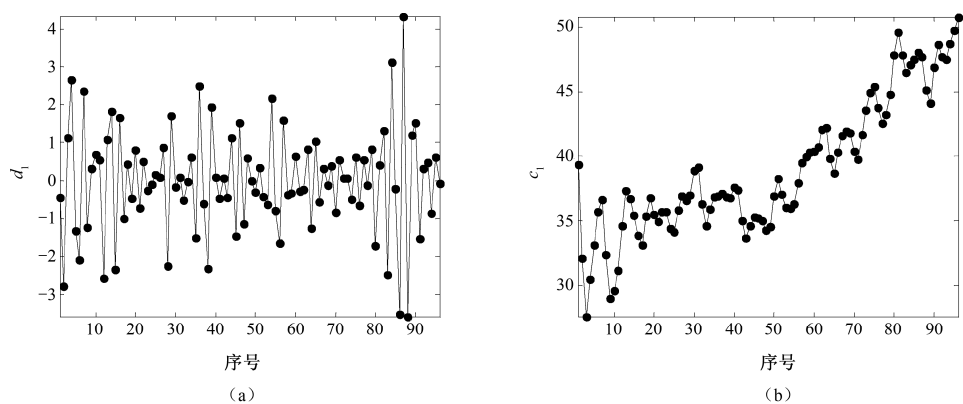


图 7-19 备件消耗数据 1 层小波变换结果

3 层小波分解与重构结果如图 7-20 所示。

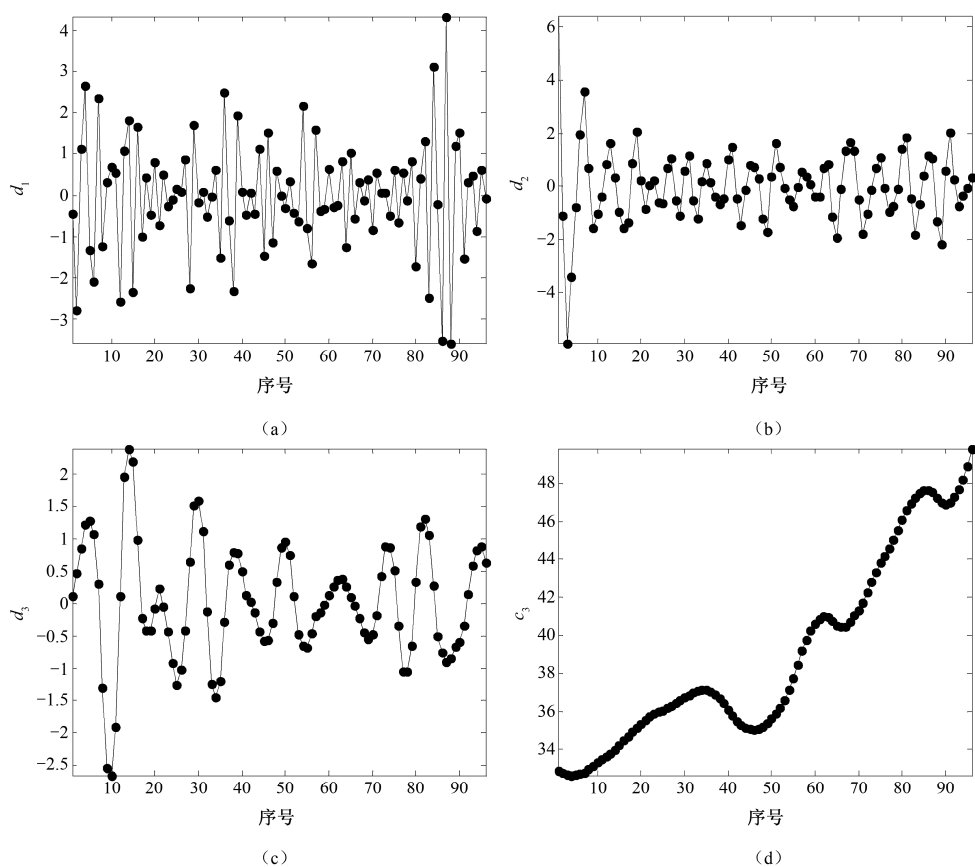


图 7-20 备件消耗数据 3 层小波变换结果

5 层小波分解与重构结果如图 7-21 所示。

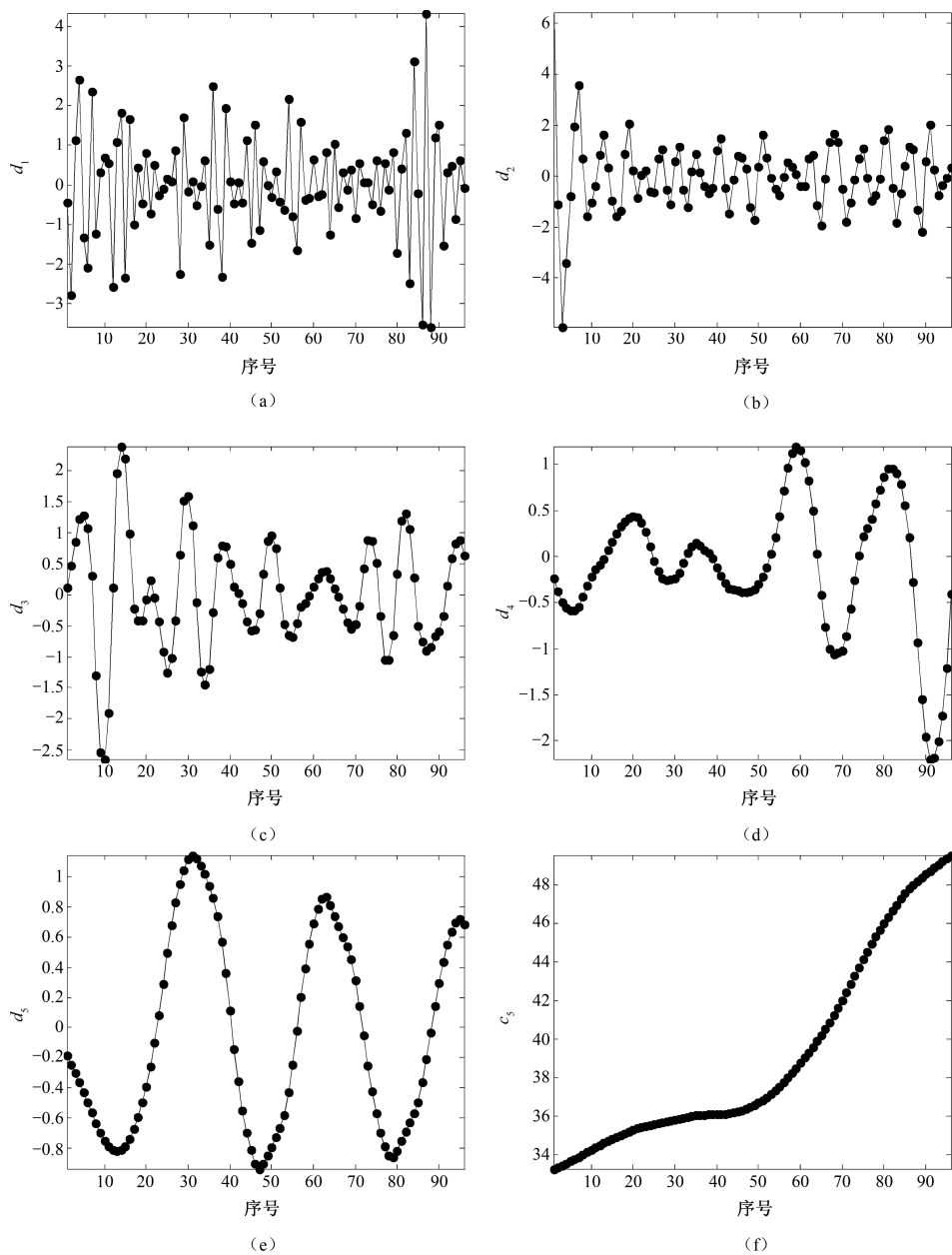


图 7-21 备件消耗数据 5 层小波变换结果

由上述小波变换结果可以得出以下结论:

(1) 由图 7-19~图 7-21 可以看出,小波变换后趋势分量的平滑度和随机分量的平稳度都较原始数据有明显提高,这说明利用小波变换可有效降低建模数据的复杂度。

(2) 根据理论分析及图 7-19~图 7-21 可知,随着小波分解层次的提高,序列频率段划分得越细,趋势分量和随机分量的平滑性和平稳性越好。但是由于在分解过程及各分量建模预测过程中必然存在计算上的误差,因此,分解层数越多,误差也会越大。在选择分解层数时,不宜过少,也不宜过多;如果待预测的时间序列数据量不是很大,则分解层数一般可以选择 3~5。本节试验中选用 Daubechies5 小波对建模数据进行 5 级分解。

3. 预测结果

采用粒子群优化算法选取模型参数。粒子群规模设为 50,混沌优化生成的粒子数目设为 30,最大迭代次数设为 50,嵌入维数和核参数的取值区间分别为 [1,20]和[0.001,1000]。作为对比试验,分别采用支持向量机回归预测模型和相关向量机回归预测模型进行预测。其中,支持向量机损失函数参数和惩罚因子的取值区间分别为[0.0001,0.1]和[1,1000]。各方法的预测值如图 7-22 所示,预测误差见表 7-9。

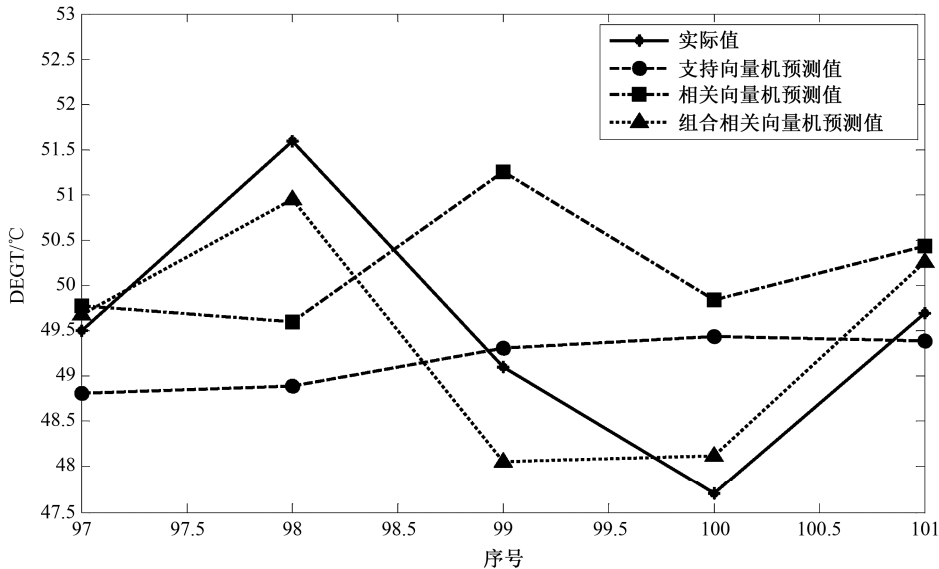


图 7-22 备件消耗预测值

表 7-9 备件消耗预测误差

序号	支持向量机		相关向量机		组合相关向量机	
	绝对误差/℃	相对误差/%	绝对误差/℃	相对误差/%	绝对误差/℃	相对误差/%
97	0.6848	1.3834	0.2913	0.5884	0.1835	0.3707
98	2.7031	5.2386	1.4619	2.8332	0.6510	1.2615
99	0.2091	0.4258	1.7232	3.5096	1.0384	2.1149
100	1.7377	3.6429	2.6049	5.4611	0.4251	0.8912
101	0.3124	0.6286	0.1919	0.3862	0.5600	1.1267
最大值	2.7031	5.2386	2.6049	5.4611	1.0384	2.1149
平均值	1.1294	2.2639	1.2547	2.5557	0.5716	1.1530

根据预测均值和预测方差计算预测值分布的概率密度,如图 7-23 和图 7-24 所示。

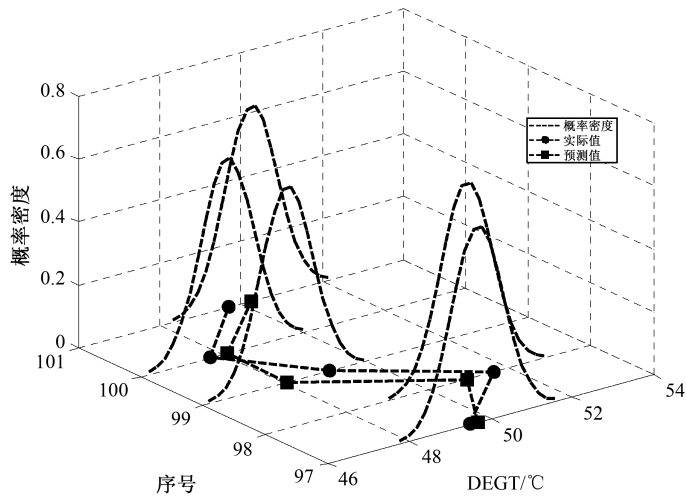


图 7-23 备件消耗预测值的概率密度

根据式 (7-85) 可以计算各预测值 90%、95%和 99%的置信区间上界,结果如图 7-24 所示。

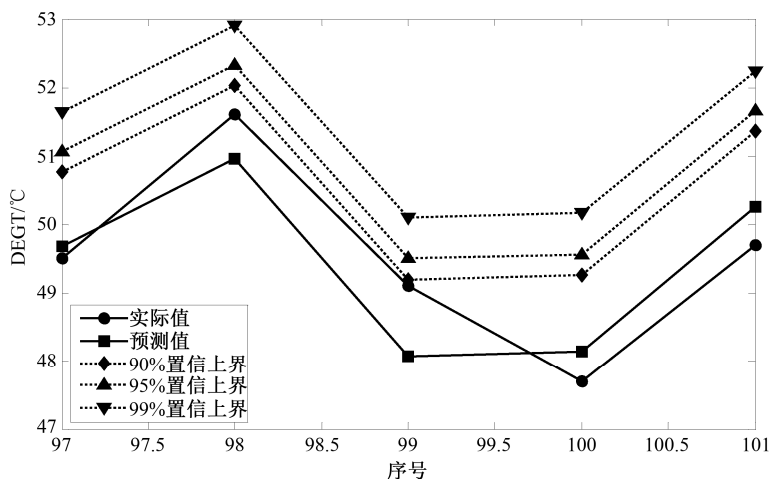


图 7-24 备件消耗预测值的置信上界

4. 结果分析

(1) 从图 7-22 和表 7-9 可以看出：相关向量机的预测性能与支持向量机相当；相对于支持向量机和相关向量机，组合相关向量机具有更高的预测精度，能够更加准确地反映备件消耗的变化趋势，备件消耗平均相对预测误差分别降低了约 1.1%和 1.4%。

(2) 从图 7-23 可以看出，本书方法预测的概率密度曲线比较紧致，说明预测的不确定性较小；备件消耗预测时均出现了预测均值小于实际值的情况，这有可能导致根据确切的预测值(预测均值)判断下一预测时刻点没有达到补充阈值，不需要对备件进行补充，然而下一预测时刻点的实际值已经超过补充阈值，从而导致备件不足，使得维修保障存在较大的缺件风险。

(3) 从图 7-24 可以看出，备件消耗的实际值均包含在置信度 90%的置信区间内。因此，结合预测值的置信区间对备件进行供应决策，得出的结果更加可靠。

基于故障分析的武器装备备件需求预测方法



随着新型武器装备的大量列装，我军的作战能力得到大大提升。这些新型武器装备大量运用现代新技术、新工艺，其系统结构复杂，电路单元数量多，元件密集度大，集成度高，保障任务重，而且有些型号武器装备数量少，因此，其备件需求预测成为一个很大的难题。不过，对于有些武器装备来说，其内部空间相对开放，工作环境相对固定，使用频率相对较高，为状态监测和故障预测创造了一定的条件。为此，本书提出对这些武器装备中具有状态监测条件的关重件的工作状态进行数据采集、处理和分析，通过故障的预测，实现备件需求的预测，及时进行备件供应和维修，从而实现精确、及时、高效的保障。

8.1 故障预测的基本原理分析

故障预测实现的前提是基于有些故障不会瞬时发生，而是要经过一段时间，即通常情况下，备件由正常状态到故障状态的转化过程不是孤立的突变，而是存在一个发展演变过程，且在故障发生前会产生一些征兆。如图 8-1 所示，从备件出现故障征兆到发展为故障（ F 点）过程中，存在一个最早的能够检测到故障征兆点，称为 P 点。在此点之前，虽然故障征兆已经出现，但是由于技术等条件

限制,不能被检测到,只有在 P 点之后,才能够通过技术手段检测到反映故障过程的特征量。

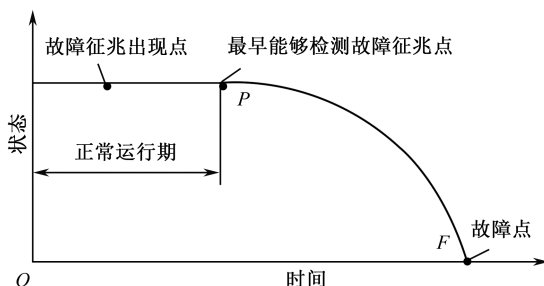


图 8-1 故障发展演变

故障预测的基本原理：通过各种状态监测手段发现故障征兆，并提取其特征信息进行数据分析，通过对可表征备件运行状态的各类状态参数的有效监测及其特征参数的分析，来反映备件运行的状态，及时跟踪故障发展的趋势，并预计故障发生时间。

将武器装备备件的运行环境和运行状态变化规律所涉及的所有现象（包括内在现象和表征现象）统称为现象空间（Phenomenon Space, PS）。利用传感器、建模仿真或其他方式获得 PS 中的元素的数值，这个数值的集合称为 PS 的测量数据集，其全体数值构成测量空间（Measurement Space, MS）。将测量空间中各元素抽象处理为能体现备件故障规律特征空间（Characteristic Space, CS）。根据特征空间和测量空间的环境参数集、传感器特性参数集等获得的结果称为预测状态空间（Forecast State Space, FSS）。下面用集合论及空间描述故障预测问题中的数据：

Φ_{WC} ：PS \rightarrow CS，表示现象空间与特征空间的映射。

Φ_{MC} ：MS \rightarrow CS，表示测量空间与特征空间的映射。

Φ_{CP} ：CS \rightarrow FSS，表示特征空间与预测状态空间的映射。

Φ_{MP} ：MS \rightarrow FSS，表示测量空间与预测状态空间的映射。

各类故障预测信息空间的表示及其之间的映射关系如图 8-2 所示。只有了解各个空间之间的映射关系，才能根据映射关系进行计算，预测出备件未来的运行状态。

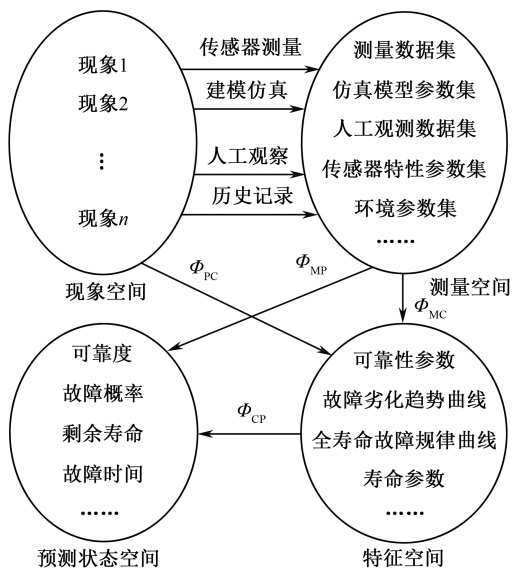


图 8-2 故障预测信息空间的表示及其联系

8.2 基于故障预测的备件需求预测系统构建

根据故障预测的基本原理,对基于故障预测的备件需求预测系统进行设计和规划,如图 8-3 所示。为提高预测的可靠性,在状态参数采集过程中,系统采用多个传感器对同一状态参数进行监测。数据预处理过程主要包括特征参数提取、剔除离群点、剔除离群数据序列和数据融合四个过程。故障预测过程是在对特征序列进行规律分析的基础上,发掘其内在规律,通过演绎推理和模型求解预测监测对象未来运行状态和故障发生时刻。经过对预测结果进行检验和评估便得到该备件需求的发生时间,进而预测未来一段时间内的备件需求量。整个系统实现的难点与重点是数据预处理和故障预测,本节分别对这两个过程进行详细论述。

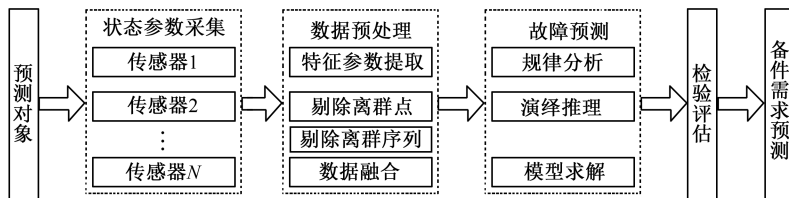


图 8-3 基于故障预测的备件需求预测系统结构

8.2.1 工作流程

基于故障预测的备件需求预测实现步骤主要包括预测对象分析、监测数据采集、数据预处理、预测方法选择、计算和预测结果分析与评价、分析决策等过程，如图 8-4 所示。具体的工作流程：首先要对预测对象进行分析，确定可表征预测对象运行状态的参数，并合理选择状态参数的监测点；通过传感器对各类状态参数数据进行采集，提取其特征信息并与数据库中存储的相关信息信息进行模糊综合匹配，形成对预测目标的状态（健康状况）评估，当无异常征兆时，则继续进行状态参数监测和采集，如果发现预测目标出现异常征兆，则根据预测对象的特点，选取合理的故障预测方法，对监测数据进行分析 and 推理，得到故障预测结果；得到预测结果后还需要对这个结果进行检验，看预测结果是否有效，当预测误差较大时应重新选择预测方法，当预测结果有效时，便可得到需要备件的具体时间。

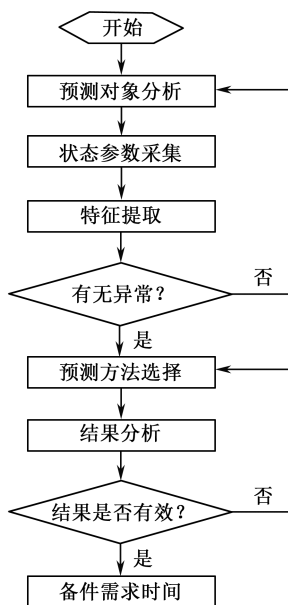


图 8-4 基于故障预测的备件需求预测系统工作流程

8.2.2 数据预处理

数据预处理的主要功能包括特征参数提取、离群点剔除、离群序列剔除和数据融合。对采集的状态参数进行特征参数提取的目的是，用大大少于原始状态参

数数目的特征信息来充分描述备件的运行状态,降低故障预测的复杂程度。离群点剔除是要剔除由于系统外部干扰或其他因素造成的远离序列一般水平的监测点。离群序列剔除是要剔除由于传感器故障产生的错误的状态参数数据序列。数据融合主要是最大限度地利用监测数据信息,提高故障预测精度。

1. 特征提取

主元分析方法常用于含有噪声和高度相关的测量数据分析,具有结构简单、便于实现的特点。采用这种方法对状态参数信息进行特征参数提取,具体步骤如下:

设采集到 n 个样本,每个样本测得 m 项指标,则样本信息数矩阵为

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \quad (8-1)$$

式中: x_{ij} 为第 i 个样本在第 j 项指标的检测值。

记

$$\left\{ \begin{array}{l} \bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \\ s_{ij} = \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) \\ v_{ij} = s_{ij} / n \\ S = (s_{ij})_{m \times n} \\ V = (v_{ij})_{m \times n} \end{array} \right. \quad (i, j = 1, 2, \cdots, m) \quad (8-2)$$

求出矩阵 V 的特征根 $\lambda_i (i = 1, 2, \cdots, m)$ 及其特征向量 $\mu_{(i)} = (u_{i1}, u_{i2}, \cdots, u_{im})$ 。令

$$z_i = \mu_{(i)} X \quad (i = 1, 2, \cdots, m) \quad (8-3)$$

$Z = (z_1, z_2, \cdots, z_m)^T$ 为样本主成分向量, z_i 为第 i 个主成分。用

$$\eta(k) = \sum_{i=1}^k \eta_i = \sum_{i=1}^k \lambda_i / \sum_{i=1}^m \lambda_i$$

计算出特征值的累计贡献率。为保证诊断的精确度,选取 $\eta(k) \geq 90\%$, 则 $z_1, z_2, \cdots, z_k (k \leq m)$ 为提取的故障特征参数。

2. 离群点剔除

采用将序列值与平滑值进行比较,检测其是否显著地大(或小)的方法来剔除离群点。这种方法是假定正常的序列值是平滑的,而离群点是突变的。用 $\overline{x(k)}$ 表示 k 时刻监测数据 $x(k)$ 的平滑值, $\overline{x(k)}^2$ 表示先对序列进行平滑再平方得到的数值, $\overline{x(k)^2}$ 表示先对序列取平方再做平滑得到的数值, $s(k)^2$ 表示样本方差, $s(k)$ 为标准差。则

$$\begin{cases} \overline{x(k)} = \alpha x(k) + (1 - \alpha) \overline{x(k-1)} \\ s(k)^2 = \overline{x(k)^2} - \overline{x(k)}^2 \end{cases} \quad (8-4)$$

如果

$$\overline{x(k)} - ns(k) < x(k+1) < \overline{x(k)} + ns(k) \quad (8-5)$$

则认为 $x(k+1)$ 正常;否则,认为是一个离群点。根据经验, α 一般取值为 $0.01 \sim 0.3$, n 一般取值为 $3 \sim 9$ 。如果 $x(k+1)$ 是一个离群点,则可用 $\hat{x}(k+1)$ 来替代,即

$$\hat{x}(k+1) = 2x(k) - x(k-1) \quad (8-6)$$

3. 离群数据序列剔除

采用关联度的方法来剔除离群数据序列。假设用三个同类型传感器采集某监测目标的状态信息,得到状态监测数据如表 8-1 所列。

表 8-1 状态监测数据

序列编号	t_0	t_1	t_2	t_3	t_4	...	t_n
序列 1	$X_{1,0}$	$X_{1,1}$	$X_{1,2}$	$X_{1,3}$	$X_{1,4}$...	$X_{1,n}$
序列 2	$X_{2,0}$	$X_{2,1}$	$X_{2,2}$	$X_{2,3}$	$X_{2,4}$...	$X_{2,n}$
序列 3	$X_{3,0}$	$X_{3,1}$	$X_{3,2}$	$X_{3,3}$	$X_{3,4}$...	$X_{3,n}$

表 8-1 中, $X_{1,0}$ 表示第一个传感器在 t_0 时刻的采样数据,其他数据类推。以序列 2 为参考序列,计算序列 1 和序列 3 相对于序列 2 的关联度。首先计算在第 k 点的关联系数:

$$\xi_{i-2}(k) = \frac{\min_i \min_k |X_{i,k} - X_{2,k}| + \rho \max_i \max_k |X_{i,k} - X_{2,k}|}{|X_{i,k} - X_{2,k}| + \rho \max_i \max_k |X_{i,k} - X_{2,k}|} \quad (i=1,3; k=0,1,2,\dots,n) \quad (8-7)$$

式中: $|X_{1,k} - X_{2,k}|$ 为数据序列 1 与序列 2 在第 k 点的绝对差; $\max_k |X_{1,k} - X_{2,k}|$ 为第一级最小差, 是序列 1 与序列 2 绝对差中的最小值; $\min_i \min_k |X_{i,k} - X_{2,k}|$ 为第二级最小差, 是所有数据序列与序列 2 的第一级最小差的最小值; $\max_i \max_k |X_{1,k} - X_{2,k}|$ 为第二级最大差, 意义与第二级最小差相似; ρ 为分辨系数, 一般在 0~1 范围取值。序列 1 相对于序列 2 的关联度为

$$r_{1 \sim 2} = \frac{1}{n} \sum_{k=1}^n \xi_{1 \sim 2}(k) \quad (8-8)$$

序列 3 与序列 2 的关联度计算类似, 另外也可以将序列 1 和序列 3 作为参考序列计算关联度。如果一个序列 i 相对于其他序列的关联度的和 r_i 小于规定值 R_s 时, 就认为这个序列为一个离群数列, 将其剔除。

4. 数据融合

如果序列 1、序列 2 和序列 3 都相关, 那么在同一个采样时刻, 表 8-1 中的值都是监测目标的状态采样值。假设在 k 时刻监测目标状态参数的真值为 $x(k)$, 采用加权自适应融合算法对采样值进行数据融合, 计算监测目标状态参数的估计值 $x'(k)$, 有

$$x'(k) = \sum_{i=1}^3 W_i x_i(k), \quad \sum_{i=1}^3 W_i = 1, \quad x_i(k) = X_{i,k} \quad (8-9)$$

由于关联度可以定量比较数据序列之间的关联程度, 即数据序列所拟合出的曲线的相似程度。采用关联度的和作为加权因子, 自动调整权值。设序列 1、序列 2 和序列 3 的关联度的和分别为 r_1 、 r_2 和 r_3 , 则对应序列的加权因子为

$$W_i = \frac{r_i}{\sum_{i=1,2,3} r_i} \quad (8-10)$$

8.2.3 故障预测

故障预测对制订武器装备修保障计划, 降低维修保障费用, 提高战备完好率和任务成功率及避免因故障造成巨大损失等方面具有重大意义。新型武器装备系统组成复杂, 结构关系模糊, 特征参数获取往往是不完整和不确定的, 这成为故

障预测实现的难点。故障预测有很多方法和模型,这些不同的故障预测方法从理论上说没有优点、缺点之分,在实际的故障预测实践中,要根据预测对象的特点和故障机理合理选择。考虑在武器装备状态监测过程中,往往存在状态监测数据获取困难、信息量少且信息不完全的实际情况,而灰色预测法在“少数据”“贫信息”和“不确定性”的问题预测中具有一定的优越性。因此,本书重点对灰色预测模型进行了深入阐述,在对同类对象故障信息进行融合的基础上,分别针对单因素和多因素两种情况建立武器装备故障预测模型。

8.2.4 备件需求确定

备件的需求来源于装备故障。根据武器装备维修保障体制,对于基层级单位来说,装备分系统或零部件出现故障,就需要进行换件维修,也就意味着发生备件需求。这里不考虑同一装备分系统或零部件多故障同时并发的情况。

根据故障预测结果,分析故障发展趋势,预测出装备分系统或零部件未来发生故障的时间,进而可以确定备件需求的发生时间。因此,对于多台装备来说,通过对某一型分系统或零部件的状态监测和故障预测,可以推算出未来一段时间内发生的故障数目,从而确定备件需求量。

8.3 基于信息融合和 IMUGM(1,1)故障预测的备件需求预测方法

在武器装备故障预测实践中,由于各种原因常出现状态监测原始数据不完整的情形,致使状态监测数据属于不等间隔期的时间序列。而常见的 GM(1,1)模型及其改进型大都是基于等间距建立的,这就大大限制了 GM(1,1)模型的实用性,且等间距是不等间距的一个特例,因此,建立 UGM(1,1)模型具有重要的现实意义。与 GM(1,1)模型一样,建模数据的背景值改造和初始值的选取也是影响 UGM(1,1)模型建模精度的主要因素。在实际工作中,随着时间的推移,旧数据的信息意义将逐渐降低,而新数据则包含大量系统运行的趋势信息。因此,及时将新数据置入模型并相应去除旧数据,所建立的模型才能反映系统当前的特征和运动趋势,符合故障预测的实际。因此,本书将研究 UGM(1,1)模型建模原理和

改进方法，并将其应用于武器装备备件需求预测。

当前的故障预测研究还面临另外一个问题，就是如何充分利用研究对象包含的丰富信息，特别是同类对象性能退化轨迹之间存在几乎相同或相近的特征。通过研究特定个体退化轨迹与同类对象各退化轨迹的相似度，有助于根据已有同类对象的退化轨迹预测特定个体未来的退化轨迹。如果对同类对象的退化轨迹数据进行相似性加权，并作为特定个体退化轨迹的输入进行故障预测，则可获得更为可信的结果。

为充分利用同类对象的故障信息，提高灰色模型的预测精度，本书提出一种基于信息融合和改进的新陈代谢不等间距灰色预测模型（Improved Metabolism Unequal Interval Grey Model, IMUGM(1,1)）相结合的预测方法，以便更加准确地预测备件需求。

8.3.1 问题描述

假设装备在恒定应力水平下工作，装备的某项性能参数 y （电流、电压或者功率等）会随着运行时间 t 呈单调递增或者递减的趋势变化。当 y 达到失效阈值 η 时，装备某系统或部件发生失效。已知 $d(d \in N)$ 个同类型系统或部件的性能退化数据，记第 $i(i=1,2,\dots,d)$ 个系统或部件的第 j 次测量时刻为 t_{ij} ，对应的性能退化测量值为 y_{ij} ，共测量 $n_i(n_i \in N)$ 次，第 i 个系统或部件的性能退化数据为

$$D_i = \{(t_{ij}, y_{ij}) | j=1, 2, \dots, n_i\}$$

记系统或部件中特定个体 Z 的 n_Z 次历史测量数据为

$$D_Z = \{(t_{Zj}, y_{Zj}) | j=1, 2, \dots, n_Z\}$$

基于同类对象的性能退化数据 D_i 和特定个体 Z 的历史测量数据 D_Z 建立特定个体 Z 的故障预测模型，即 y 与 t 之间的函数关系为

$$y = f_Z(t) \quad (8-11)$$

结合 η 预测 Z 的失效时间 T ，即求解 $\eta = f_Z(T)$ 从而得到剩余寿命预测值 $(T-t)$ 。根据 Z 的监测数据 (t, y_Z) ，对式（8-11）进行更新实现剩余寿命预测，最终确定备件需求。

8.3.2 IMUGM(1,1)模型的建立

通过改进初始值选取方式和背景值构造形式，及运用新陈代谢思想对样本数

据进行及时更新, 建立 IMUGM(1,1)。

1. 改进初始值的 UGM(1,1)模型

设原始序列 $X^{(0)} = \{x^{(0)}(k_1), x^{(0)}(k_2), \dots, x^{(0)}(k_n)\}$, 相邻分量之间的间距为 $\Delta k_i = k_i - k_{i-1}$ ($i = 1, 2, \dots, n$)。若 Δk_i 为常数, 则称序列 $X^{(0)}$ 为等间距序列; 若 Δk_i 不为常数, 则称 $X^{(0)}$ 为不等间距序列。

不等间距 GM(1,1)模型建模过程如下:

(1) 对 $X^{(0)}$ 做一次累加生成, 得

$$X^{(1)} = \{x^{(1)}(k_1), x^{(1)}(k_2), \dots, x^{(1)}(k_n)\}$$

式中: $x^{(1)}(k_i) = \sum_{j=1}^i x^{(0)}(k_j) \times \Delta k_j$ ($j = 2, 3, \dots, n$)。

规定 $x^{(1)}(k_1) = x^{(0)}(k_1)$ 。

(2) 由一阶生成模块 $X^{(1)}$ 建立 GM(1,1)模型, 对应的白化微分方程为

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b \quad (8-12)$$

灰色微分方程为

$$x^{(0)}(k_{i+1}) \times \Delta k_{i+1} + az^{(1)}(k_{i+1}) = b\Delta k_{i+1} \quad (8-13)$$

式中: $z^{(1)}(k_{i+1})$ 为 $x^{(1)}(t)$ 在 $[k_i, k_{i+1}]$ 上的背景值, 且有

$$z^{(1)}(k_{i+1}) = \frac{1}{2}(x^{(1)}(k_{i+1}) + x^{(1)}(k_i)) \quad (8-14)$$

(3) 灰色微分方程的最小二乘参数估计为

$$\hat{\alpha} = (\hat{a}, \hat{b})^T = (B^T B)^{-1} B^T Y \quad (8-15)$$

式中

$$B = \begin{bmatrix} -z^{(1)}(k_2) & \Delta k_2 \\ -z^{(1)}(k_3) & \Delta k_3 \\ \vdots & \vdots \\ -z^{(1)}(k_n) & \Delta k_n \end{bmatrix}, \quad Y = \begin{bmatrix} x^{(0)}(k_2) \times \Delta k_2 \\ x^{(0)}(k_3) \times \Delta k_3 \\ \vdots \\ x^{(0)}(k_n) \times \Delta k_n \end{bmatrix}$$

(4) 方程式 (8-12) 的离散解, 即时间响应式为

$$\hat{x}^{(1)}(k_i) = \left[x^{(1)}(k_1) + c - \frac{b}{a} \right] e^{-a(k_i - k_1)} + \frac{b}{a} \quad (8-16)$$

式中: c 为待定初始值优化量。

(5) 还原到原始数据为

$$\hat{x}^{(0)}(k_{i+1}) = \frac{1}{k_{i+1} - k_i} \left(1 - e^{a(k_{i+1} - k_i)} \right) \left[x^{(0)}(k_i) + c - \frac{b}{a} \right] e^{-a(k_{i+1} - k_i)} \quad (8-17)$$

2. 改进背景值的 UGM(1,1)模型

从上面建模过程可知,模型的拟合和预测精度取决于系数 a 和 b ,而 a 和 b 的求解依赖于背景值 $z^{(1)}(k_{i+1})$ 的构造形式,因此,背景值 $z^{(1)}(k_{i+1})$ 的值是直接影响 UGM(1,1)模型精度和适应性的关键因素。下面针对 UGM(1,1)模型背景值的改进方法做进一步研究,以提高 UGM(1,1)模型的精度及适应范围。

将式(8-13)在区间 $[k_i, k_{i+1}]$ 上积分,有

$$x^{(1)}(k_{i+1}) - x^{(1)}(k_i) + a \int_{k_i}^{k_{i+1}} x^{(1)}(t) dt = b(k_{i+1} - k_i) \quad (i=1, 2, \dots, n-1) \quad (8-18)$$

将式(8-18)和式(8-13)比较可知,用一阶线性微分式(8-12)的解来逼近 $x^{(1)}(k)$,其误差来源于

$$z^{(1)}(k_{i+1}) = \frac{1}{2} (x^{(1)}(k_{i+1}) + x^{(1)}(k_i))$$

代替 $\int_{k_i}^{k_{i+1}} x^{(1)}(t) dt$ 所致。为消除此误差,不妨记 $Z^{(1)}(k_{i+1}) = \int_{k_i}^{k_{i+1}} x^{(1)}(t) dt$ 。

由于方程式(8-12)的解为指数形式,因此, $x^{(1)}(t)$ 可用指数曲线 $x^{(1)}(t) = Le^{Mt} + N$, L 、 M 和 N 为待定常数,其求解方法可参考等间距灰色模型建模背景值改进方法。

$$M = \frac{1}{\Delta k_{i+1}} \times \left\{ \ln [x^{(1)}(k_{i+1}) - x^{(1)}(k_i)] - \ln x^{(1)}(k_i) \right\}, \quad N = \frac{x^{(1)}(k_i) \cdot [x^{(1)}(k_i)]^{\frac{k_{i+1}}{\Delta k_{i+1}}}}{x^{(0)}(k_{i+1})^{\frac{k_i}{\Delta k_{i+1}}} - x^{(1)}(k_i)}$$

背景值为

$$z^{(1)}(k_{i+1}) = \frac{x^{(1)}(k_{i+1}) - x^{(1)}(k_i)}{\ln [x^{(1)}(k_{i+1}) - x^{(1)}(k_i)] - \ln x^{(1)}(k_i)} \times \Delta k_{i+1} + \frac{x^{(1)}(k_i) \cdot [x^{(1)}(k_i)]^{\frac{k_{i+1}}{\Delta k_{i+1}}}}{x^{(0)}(k_{i+1})^{\frac{k_i}{\Delta k_{i+1}}} - x^{(1)}(k_i)} \quad (8-19)$$

用式(8-19)代替式(8-14)的背景值构造公式,其他建模过程同改进初始值的 UGM(1,1)模型。

3. 基于新陈代谢思想优化的 UGM(1,1)模型

添加 $\hat{x}^{(0)}(k_{i+1})$ 到原始数据序列并删除 $\hat{x}^{(0)}(k_i)$,经上述运算得 $\hat{x}^{(0)}(k_{i+2})$,依

此类推直到最后求得 $\hat{x}^{(0)}(k_n)$ 。

8.3.3 基于信息融合和 IMUGM(1,1)故障预测的备件需求预测模型构建

本节所提出的备件需求确定方法分为四个阶段：①建模阶段，对同类对象和特定个体进行基于 IMUGM(1,1)的故障预测建模；②信息融合阶段，为了充分利用研究对象的故障信息，运用加权思想确定特定个体 Z 对同类对象和特定个体 Z 故障预测模型的隶属度权值；③预测阶段，基于隶属度权值和同类对象、特定个体 Z 的故障预测模型建立特定个体 Z 的最终故障预测模型，然后结合监测数据更新隶属度和特定个体最终故障预测模型，实现故障预测；④备件需求确定阶段，根据故障预测结果，结合装备使用和保障情况，确定备件需求。具体过程如图 8-5 所示。

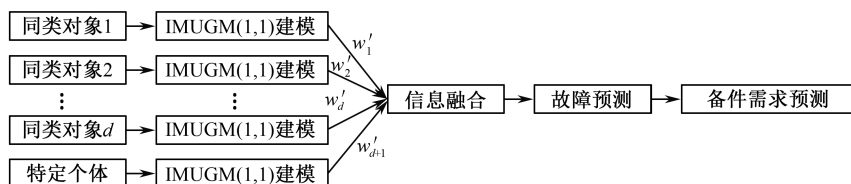


图 8-5 基于信息融合和 IMUGM(1,1)故障预测的备件需求预测框架图

1. 故障预测建模

首先根据 8.3.2 节相关原理，先运用同类对象的历史监测值 $y_{ij}(i=1,2,\cdots,d; j=1,2,\cdots,n_i)$ ，建立 d 个基于 IMUGM(1,1)的故障预测模型；再运用特定个体 Z 的历史监测值 y_{Zj} ，建立第 $d+1$ 个基于 IMUGM(1,1)的故障预测模型；统一记为 $\{f_i(t)|i=1,2,\cdots,d+1\}$ 。

2. 故障信息融合及备件需求确定

将特定个体 Z 的历史测量时刻向量 t_Z 分别代入 $d+1$ 个故障预测模型，计算出 $d+1$ 个预测值或拟合值向量 $Y_i(i=1,2,\cdots,d+1)$ ；接着计算特定个体 Z 的测量值向量 Y_Z 与 Y_i 的欧几里得 (Euclid) 距离 $d_{Zi}(i=1,2,\cdots,d,d+1)$ ；然后根据 d_{Zi} 确定出 Y_Z 对同类对象和特定个体 Z 的故障预测模型的隶属度权值 $w'_i(i=1,2,\cdots,d,d+1)$ ，对 $d+1$ 个故障预测模型进行加权得到 Z 的最终故障预测

模型 $f_Z(t)$; 最后结合 Z 的实时测量数据 $[t, y_Z]$ 对 t_Z 和 Y_Z 进行扩充, 依次更新 Y_i 、 d_{Zi} 和 $w'_i (i=1, 2, \dots, d, d+1)$, 实现 $f_Z(t)$ 的实时更新, 并结合 η 进行故障预测。

隶属度加权法 (Weight Method of Merlevship Degree, WMMD) 步骤如下:

(1) 令 $s = n_Z$, $t_0 = t_{Zs}$, $t_Z = [t_{Z1}, t_{Z2}, \dots, t_{Zs}]^T$, $Y_Z = [y_{Z1}, y_{Z2}, \dots, y_{Zs}]^T$ 。

(2) 将 t_Z 代入 $\{f_i(t) | i=1, 2, \dots, d+1\}$ 算得 $Y_i = [f_i(t_{Z1}), f_i(t_{Z2}), \dots, f_i(t_{Zs})]^T$; 计算 Y_Z 与 Y_i 的欧几里得距离, 即

$$d_{Zi} = \left[\sum_{j=1}^s (y_{Zj} - f_i(t_{Zj}))^2 \right]^{1/2} \quad (i=1, 2, \dots, d+1)$$

(3) 计算 Y_Z 对特定个体 Z 和同类对象的故障预测模型的隶属度权值

$$w'_i = \left[\sum_{j=1}^n (d_{Zi} d_{Zj}^{-1})^2 \right]^{-1} \quad (i=1, 2, \dots, d+1)$$

如果存在 i , 使得 $d_{Zi} = 0$, 则有 $w'_i = 1$, $w'_j = 0 (j \neq i)$ 。

(4) 建立特定个体 Z 的最终故障预测模型

$$f_Z(t) = \sum_{i=1}^n w'_i f_i(t) (i=1, 2, \dots, d+1)$$

求解 $\eta = f_Z(T)$ 得 Z 的故障时间预测值 T , 计算剩余寿命 $(T - t_0)$ 。

(5) 获取 Z 的最新监测数据 (t, y_Z) , 令 $t_0 = t$, $s = s+1$, $t_Z = [t_z; t]$, $Y_Z = [Y_z; y_Z]$, 转至步骤 (2)。

(6) 根据备件剩余寿命预测值, 确定未来一段时期的备件需求量。

8.3.4 示例分析

电源组合是导弹制导雷达系统的重要组成部分, 用于提供雷达装备工作所需的各种稳压电源。电源组合的失效方式为退化失效, 通常测试的是电压信号, 即发现输出电压有一定的偏离但还不足以维修时, 就需要密切关注其输出电压并能进行准确的故障预测。下面以某型导弹制导雷达系统的波束控制系统中 P11 电源系统为例, 对该系统中最复杂的 $\Pi \Pi 98-3$ 电源组合采集电压信号数据, 分析电压测试信号的特点, 对其进行故障预测。设备平均每天工作 2h。

图 8-6 为 P11 系统电源组合功能框图。图 8-7 为 $\Pi \Pi 98-3$ 电源组合中的一路, 输出为 $\pm 15V$ 稳定直流电。具体工作过程: 400Hz 的 220V 交流电经电桥整流后, 送到变压器中进行降压; 然后通过全波整流和滤波变成直流; 最后通过串联负反

馈稳压电路输出 15V 直流电压，最大不超过 20V，否则电源组合会发生故障。

由于 П П 98-3 电源组合对波束控制系统的影响主要是由电源电压的退化引起的，因此电源电压的测试是一项重要内容，表 8-2 和图 8-8 给出了 3 套 П П 98-3 电源组合的部分电压监测采样值，可以看出电压值大体呈上升趋势（下降趋势相同，这里以上升趋势为例）。

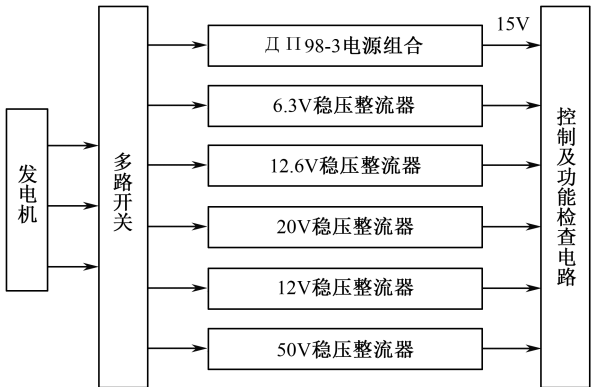


图 8-6 П П 98-3 系统电源组合功能框图

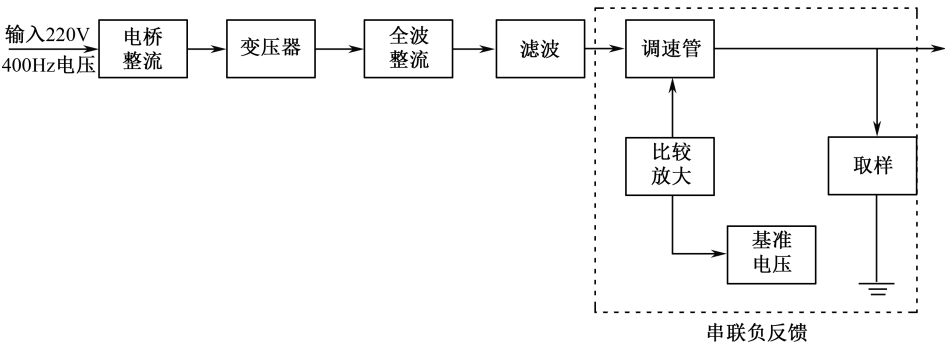


图 8-7 П П 98-3 电源组合原理框图

表 8-2 П П 98-3 电源组合的状态监测数据

时刻		200	260	320	380	440	500	550	600	650
监测值	1	15.06	15.17	15.30	15.49	15.75	16.01	16.16	16.39	16.67
	2	15.00	15.08	15.19	15.44	15.56	15.82	15.99	16.24	16.44
	3	15.02	15.11	15.23	15.40	15.63	15.85	16.06	16.27	16.53

续表

时刻		700	740	780	820	860	890	920	950	980
监测值	1	16.79	17.08	17.41	17.79	18.22	18.65	19.01	19.43	19.87
	2	16.71	16.96	17.32	17.59	17.93	18.42	18.78	19.17	19.64
	3	16.82	17.12	17.43	17.76	18.15	18.57	18.92	19.34	19.78

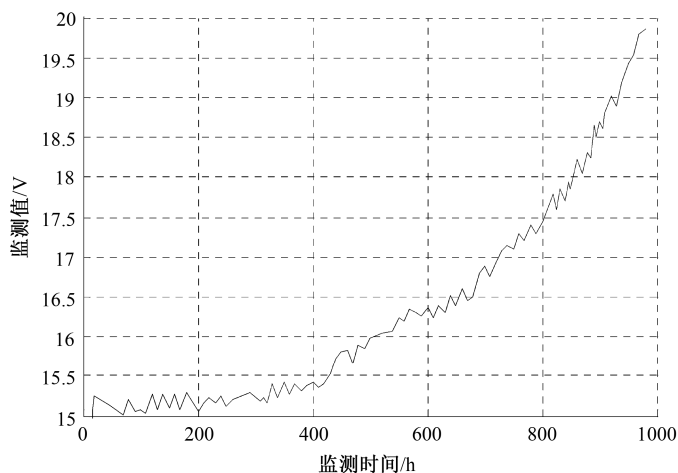


图 8-8 ДП 98-3 电源组合的电压监测值

1. 模型建立

对于同类对象 1 和 2, 从第 2 个数据开始的前 13 组数据用于构建预测模型; 对于特定个体 3, 从第 1 个数据开始的前 13 组数据用于构建预测模型, 后 5 组数据作为预测检验数据。采用基于信息融合和 IMUGM(1,1)模型相结合的方法进行故障预测, 使用 Matlab 软件进行仿真计算分析。

原始数据序列为

$$X_1^{(0)} = \{15.17, 15.30, 15.49, 15.75, 16.01, 16.16, 16.39, 16.67, 16.79, 17.08, 17.41, 17.79, 18.22\}$$

$$X_2^{(0)} = \{15.08, 15.19, 15.44, 15.56, 15.82, 15.99, 16.24, 16.44, 16.71, 16.96, 17.32, 17.59, 17.93\}$$

$$X_3^{(0)} = \{15.02, 15.11, 15.23, 15.40, 15.63, 15.85, 16.06, 16.27, 16.53, 16.82, 17.12, 17.43, 17.76\}$$

根据式 (8-12) ~ 式 (8-19), 分别建立 IMUGM(1,1)模型, 编程计算得

$$\hat{x}_1^{(0)}(t_{14}) = 18.1963, \quad \hat{x}_2^{(0)}(t_{14}) = 17.9538, \quad \hat{x}_3^{(0)}(t_{14}) = 18.0726$$

分别添加 $x_1^{(0)}(t_{15})$, $x_2^{(0)}(t_{15})$, $\hat{x}_1^{(0)}(t_{14})$, 并删除 $x_1^{(0)}(t_2)$, $x_2^{(0)}(t_2)$, $x_3^{(0)}(t_1)$, 计算得

$\hat{x}_1^{(0)}(t_{15})=18.5839, \hat{x}_2^{(0)}(t_{15})=18.3195, \hat{x}_3^{(0)}(t_{15})=18.2162$

依此类推，直到得到全部数据。

2. 故障预测

根据欧几里得距离 d_{Zi} 和隶属度权值 w'_i 的定义，可以编程计算得 $w'_{11}=0.5039, w'_{21}=0.1592, w'_{31}=0.3369$ ，进而求得 $\hat{x}^{(0)}(t_{14})=18.1160$ 。依次类推，最后求得 $w'_{15}=0.3273, w'_{25}=0.1892, w'_{35}=0.4835, \hat{x}^{(0)}(t_{18})=19.8372$ 。

同时，为检验本节预测方法的预测效果，将信息融合和 IMUGM(1,1)模型相结合的预测方法与传统 UGM(1,1)模型和 IMUGM(1,1)模型的预测数据进行比较分析，各预测方法或模型的拟合效果及预测结果分别如表 8-3、图 8-9 及表 8-4、图 8-10 所示，并分别以预测值（Forecast Value，FV）、相对误差（RPE）和平均相对误差（MRE）表示。

表 8-3 不同方法的拟合值及相对误差

监测数据		传统 UGM(1,1)模型		IMUGM(1,1)模型		本节预测方法	
工作时间/h	电压值/V	拟合值	相对误差/%	拟合值	相对误差/%	拟合值	相对误差/%
200	15.02	15.0200	0	15.0200	0	15.0200	0
260	15.11	14.7395	2.45	15.0568	0.35	15.0768	0.22
320	15.23	14.9113	2.09	15.1261	0.68	15.1782	0.34
380	15.40	15.1758	1.46	15.2905	0.71	15.3557	0.29
440	15.63	15.5271	0.65	15.4937	0.87	15.5916	0.25
500	15.85	15.9996	0.94	15.8959	0.29	15.8239	0.16
550	16.06	16.3428	1.76	16.1892	0.80	16.1093	0.31
600	16.27	16.7735	3.09	16.3615	0.56	16.3529	0.51
650	16.53	16.9344	2.44	16.4659	0.39	16.5674	0.23
700	16.82	17.0686	1.48	16.6437	1.05	16.7588	0.36
740	17.12	17.2952	1.02	16.9586	0.94	17.0419	0.45
780	17.43	17.4829	0.25	17.3324	0.45	17.3868	0.24
820	17.76	17.8164	0.32	17.7067	0.30	17.7324	0.16
平均相对误差/%		1.94		0.56		0.27	

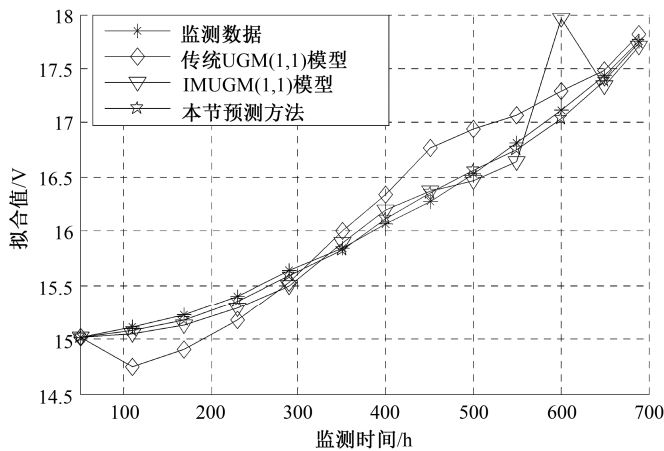


图 8-9 不同方法的拟合结果

表 8-4 不同方法的预测值及相对误差

监测数据		传统 UGM(1,1) 模型		IMUGM(1,1) 模型		本节预测方法	
工作时间/h	电压值/V	预测值	相对误差/%	预测值	相对误差/%	预测值	相对误差/%
860	18.15	17.5792	3.14	18.0126	0.76	18.1160	0.19
890	18.57	18.1295	2.37	18.3294	1.30	18.4879	0.44
920	18.92	18.5683	1.86	18.9943	0.39	18.8135	0.56
950	19.34	18.9148	2.20	19.5379	1.02	19.4594	0.62
980	19.78	19.3061	2.39	19.9341	0.78	19.8372	0.29
平均相对误差/%		2.39		0.85		0.42	

3. 结果分析及应用

从表 8-3、图 8-9 及表 8-4、图 8-10 中可以看出, 本节提出的预测方法拟合效果最好、预测精度最高, 其次是 IMUGM(1,1)模型, 最差的是传统 UGM(1,1)模型。一方面 IMUGM(1,1)模型采用改进初始值选取方法、改进背景值构造形式及新陈代谢思想, 对建模过程进行了改进和优化, 使得预测精度有了大幅度提高; 另一方面在 IMUGM(1,1)模型的基础上, 通过信息融合技术充分利用了同类对象的故障信息, 使预测结果的相对误差减小了 1/2, 说明该方法的预测效果是明显的。

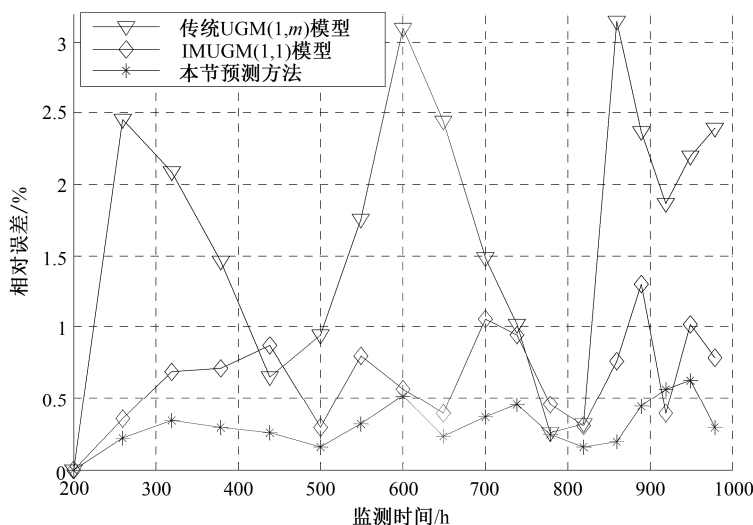


图 8-10 相对误差对比

以第 13 个工作时间所对应的时间点作为参照，预测下一个备件需求发生的时间。根据上面建立的故障预测模型，可计算出当电压值 20V 时，工作时间为 1003h，以每天工作 2h 进行折算，则有 $(1003-820)/2=92$ ，即在未来 92 天左右时需要 1 个备件。

8.4 基于信息融合和 IMUGM(1,m,w)故障预测的备件需求预测方法

对于武器装备来说，影响故障发展的因素较多，且各因素之间一般具有内在的联系，有时几个指标或参数一起才能反映装备的状态。因此，很多情况下是一种多因素情况下的故障预测。另外，传统灰色模型的背景值构造采用基本的紧邻均值法，这种背景值构造方式不足以反映背景值在建模中的作用，且当序列变化急剧时，预测误差较大。

为充分利用同类对象的故障信息，提高灰色模型的预测精度，本书提出一种基于信息融合和改进的多变量新陈代谢不等间距加权灰色预测模型（Improved multi-variables Metabolism Unequal Interval weighted Grey Model，IMUGM(1,m,w)）相结合的预测方法，以便更加准确地预测备件需求。

8.4.1 问题描述

假设装备在恒定应力水平下工作, 装备的某项性能参数 y (电流、电压或者功率等) 会随着运行时间 t 呈单调递增或者递减的趋势变化, 当 y 达到失效阈值 η 时装备某系统或部件发生失效。已知 $d(d \in N)$ 个同类型系统或部件的性能退化数据, 记第 $i(i=1,2,\dots,d)$ 个系统或部件的第 j 次测量时刻为 t_{ij} , 对应的性能退化测量值为 y_{ij} , 共测量 $n_i(n_i \in N)$ 次, 第 i 个系统或部件的性能退化数据为

$$D_i = \{(t_{ij}, y_{ij}) | j=1, 2, \dots, n_i\}$$

记系统或部件中特定个体 Z 的 n_Z 次历史测量数据为

$$D_Z = \{(t_{Zj}, y_{Zj}) | j=1, 2, \dots, n_Z\}$$

基于同类对象的性能退化数据 D_i 和特定个体 Z 的历史测量数据 D_Z 建立特定个体 Z 的故障预测模型, 即 y 与 t 之间的函数关系为

$$y = f_Z(t) \quad (8-20)$$

结合 η 预测 Z 的失效时间 T , 即求解 $\eta = f_Z(T)$ 从而得到剩余寿命预测值 $(T-t)$ 。根据 Z 的监测数据 (t, y_Z) , 对式 (8-20) 进行更新实现剩余寿命预测, 最终确定备件需求。

8.4.2 IMUGM(1,m,w)模型的建立

令 $x_i^{(0)}(t)$ ($i=1,2,\dots,m$) 为 m 个灰时间序列, $x_i^{(0)}(t_j)$ ($j=1,2,\dots,n$) 为各因素在时刻 t_j 的观测值, 原始数据为

$$\mathbf{X}^{(0)} = \{X^{(0)}(t_1), X^{(0)}(t_2), \dots, X^{(0)}(t_n)\} = \begin{bmatrix} x_1^{(0)}(t_1) & x_1^{(0)}(t_2) & \dots & x_1^{(0)}(t_n) \\ x_2^{(0)}(t_1) & x_2^{(0)}(t_2) & \dots & x_2^{(0)}(t_n) \\ \vdots & \vdots & & \vdots \\ x_m^{(0)}(t_1) & x_m^{(0)}(t_2) & \dots & x_m^{(0)}(t_n) \end{bmatrix} \quad (8-21)$$

将间距作为乘子, 对原始数据列进行一次累加生成, 得新数据矩阵

$$\mathbf{X}^{(1)} = \{X^{(1)}(t_1), X^{(1)}(t_2), \dots, X^{(1)}(t_n)\} = \begin{bmatrix} x_1^{(1)}(t_1) & x_1^{(1)}(t_2) & \dots & x_1^{(1)}(t_n) \\ x_2^{(1)}(t_1) & x_2^{(1)}(t_2) & \dots & x_2^{(1)}(t_n) \\ \vdots & \vdots & & \vdots \\ x_m^{(1)}(t_1) & x_m^{(1)}(t_2) & \dots & x_m^{(1)}(t_n) \end{bmatrix} \quad (8-22)$$

式中:

$$x_i^{(1)}(t_j) = \sum_{k=1}^j x_i^{(0)}(t_k) \cdot \Delta t_k \quad (\Delta t_k = t_k - t_{k-1} \neq \text{const})$$

对 $x_i^{(1)}(t_j)$ 建立白化形式的微分方程组:

$$\begin{cases} \frac{dx_1^{(1)}}{dt_1} = a_{11}x_1^{(1)}(t_1) + a_{12}x_2^{(1)}(t_2) + \cdots + a_{1m}x_m^{(1)}(t_m) + b_1 \\ \frac{dx_2^{(1)}}{dt_2} = a_{21}x_1^{(1)}(t_1) + a_{22}x_2^{(1)}(t_2) + \cdots + a_{2m}x_m^{(1)}(t_m) + b_2 \\ \vdots \\ \frac{dx_m^{(1)}}{dt_m} = a_{m1}x_1^{(1)}(t_1) + a_{m2}x_2^{(1)}(t_2) + \cdots + a_{mm}x_m^{(1)}(t_m) + b_m \end{cases} \quad (8-23)$$

$$\text{记 } \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}, \quad \mathbf{B} = (b_1, b_2, \dots, b_m)^T, \quad \text{则式 (8-23) 的向量表达式为}$$

$$\begin{cases} \frac{dx_i^{(1)}(t_j)}{dt_j} + a_i x_i^{(1)}(t_j) = b_i \\ x_i^{(1)}(t_1) = x_i^{(0)}(t_1) \end{cases} \quad (8-24)$$

将上述微分方程离散化

$$x_i^{(0)}(t_{j+1}) + a_i z_i^{(1)}(t_{j+1}) = b_i \quad (i=1, 2, \dots, m) \quad (8-25)$$

式中: $z_i^{(1)}(t_{j+1})$ 为 $x_i^{(1)}(t_j)$ 在区间 $[t_j, t_{j+1}]$ 上的背景值。

在传统灰色模型中, 背景值的构造采用紧邻均值法, 即

$$z_i^{(1)}(t_{j+1}) = \frac{1}{2} [x_i^{(1)}(t_{j+1}) + x_i^{(1)}(t_j)]$$

针对这种背景值的构造方式不足以反映背景值在建模中的作用, 且当序列变化急剧时, 预测误差较大的缺陷。采用引入加权因子 w_i 的方式来优化背景值构造, 令

$$z_i^{(1)}(t_{j+1}) = w_i x_i^{(1)}(t_{j+1}) + (1 - w_i) x_i^{(1)}(t_j) \quad (8-26)$$

式中: w_i 为背景值最佳生成系数 ($0 < w_i < 1$)。它的值通过满足原始值与模型预测值的平均相对误差达到最小来确定。平均相对误差为

$$\bar{\Delta} = \frac{1}{n} \sum_{j=1}^n \frac{|x_i^{(0)}(t_j) - \hat{x}_i^{(0)}(t_j)|}{x_i^{(0)}(t_j)} \quad (8-27)$$

令 $\mathbf{a}_i = (a_{i1}, a_{i2}, \dots, a_{in}, b_i)^T$ ($i=1, 2, \dots, m$)。用矩阵形式表示式 (8-25)，即

$$\mathbf{Y}_i = \mathbf{L} \cdot \hat{\mathbf{a}}_i \quad (8-28)$$

式中

$$\mathbf{L} = \begin{bmatrix} -z_1^{(1)}(t_2) & -z_2^{(1)}(t_2) & \cdots & -z_m^{(1)}(t_2) & 1 \\ -z_1^{(1)}(t_3) & -z_2^{(1)}(t_3) & \cdots & -z_m^{(1)}(t_3) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ -z_1^{(1)}(t_n) & -z_2^{(1)}(t_n) & \cdots & -z_m^{(1)}(t_n) & 1 \end{bmatrix} \quad \mathbf{Y}_i = \{x_i^{(0)}(t_2), x_i^{(0)}(t_3), \dots, x_i^{(0)}(t_m)\}^T$$

采用最小二乘法求解式 (8-28)，可得

$$\hat{\mathbf{a}}_i = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T \mathbf{Y}_i \quad (8-29)$$

则可以得到 \mathbf{A} 和 \mathbf{B} 的辨识值：

$$\hat{\mathbf{A}} = \begin{bmatrix} \hat{a}_{11} & \hat{a}_{12} & \cdots & \hat{a}_{1m} \\ \hat{a}_{21} & \hat{a}_{22} & \cdots & \hat{a}_{2m} \\ \vdots & \vdots & & \vdots \\ \hat{a}_{m1} & \hat{a}_{m2} & \cdots & \hat{a}_{mm} \end{bmatrix}, \quad \hat{\mathbf{B}} = (\hat{b}_1, \hat{b}_2, \dots, \hat{b}_m)^T$$

将求得的灰参数 $\hat{\mathbf{a}}_i$ 代入白化微分方程式 (8-25)，得一次累加序列的表达式为

$$x_i^{(1)}(t_j) = e^{\hat{A}(t_j - t_1)} x_i^{(1)}(t_1) + \hat{A}^{-1} (e^{\hat{A}(t_j - t_1)} - \mathbf{I}) \hat{\mathbf{B}} \quad (8-30)$$

式中

$$e^{\hat{A}(t_j - t_1)} = \mathbf{I} + \hat{\mathbf{A}}(t_j - t_1) + \frac{\hat{\mathbf{A}}^2}{2!} (t_j - t_1)^2 + \cdots = \mathbf{I} + \sum_{k=1}^{\infty} \frac{\hat{\mathbf{A}}^k}{k!} (t_j - t_1)^k$$

再将上式累减还原，得原始监测数据的预测值表达式为

$$\hat{x}_i^{(0)}(t_{j+1}) = \frac{\hat{x}_i^{(1)}(t_{j+1}) - \hat{x}_i^{(1)}(t_j)}{\Delta t_{j+1}} \quad (8-31)$$

$\hat{X}^{(0)}(t_{j+1})$ 可由式 (8-31) 计算获得，然后添加 $\hat{X}^{(0)}(t_{j+1})$ 并删除 $X^{(0)}(t_1)$ ，构造新序列，即

$$\{X^{(0)}(t_2), X^{(0)}(t_3), \dots, \hat{X}^{(0)}(t_{j+1})\} \quad (8-32)$$

取 $X^{(1)}(t_2) = X^{(0)}(t_2)$ ，执行式 (8-21) ~ 式 (8-31) 的操作，得 $\hat{X}^{(0)}(t_{j+2})$ 。

依此类推，直到最后求 $\hat{X}^{(0)}(t_n)$ 。

8.4.3 基于信息融合和 IMUGM(1,m,w)故障预测的备件需求预测模型构建

本书所提出的备件需求确定方法可分为四个阶段：①建模阶段，对同类对象和特定个体进行基于 IMUGM(1,m,w)的故障预测建模；②信息融合阶段，为了充分利用研究对象的故障信息，运用加权思想确定特定个体 Z 对同类对象和特定个体 Z 故障预测模型的隶属度权值；③预测阶段，基于隶属度权值和同类对象、特定个体 Z 的故障预测模型建立特定个体 Z 的最终故障预测模型，然后结合监测数据更新隶属度和特定个体最终故障预测模型，实现故障预测；④备件需求确定阶段，根据故障预测结果，结合装备使用和保障情况，确定备件需求。具体过程如图 8-11 所示。

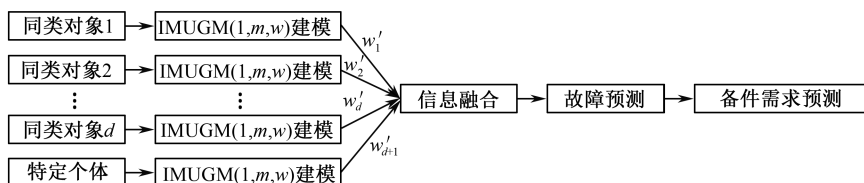


图 8-11 基于信息融合和 IMUGM(1,m,w)的备件需求预测框架图

1. 故障预测建模

首先根据 8.4.2 节所讲的原理，运用同类对象的历史监测值 $y_{ij}(i=1,2,\cdots,d; j=1,2,\cdots,n_i)$ ，建立 d 个基于 IMUGM(1,m,w)的故障预测模型；再运用特定个体 Z 的历史监测值 y_{Zj} ，建立第 $d+1$ 个基于 IMUGM(1,m,w)的故障预测模型，统一记为 $\{f_i(t)|i=1,2,\cdots,d+1\}$ 。

2. 故障信息融合及备件需求确定

应用隶属度加权法进行故障信息融合，具体过程参照 8.3.3 节内容。

8.4.4 示例分析

高频模拟器模块用来产生各种脉冲波形来模拟攻击目标的距离、方位等信

息，并回馈给导引头以达到检测目的，是主动雷达制导导弹检测设备中重要的分系统。以某型导弹检测设备为例，它的高频模拟器使用 RS-422 串口与上位主控制器通信，通过自行定义的通信协议，可控制产生相应脉冲，经与高频信号调制并放大，最后由天线发射出去。末端经过一个功率分配器，将产生的高频已调制信号分给自检接口进行设备的自检，其原理如图 8-12 所示。其中，功放组件是高频模拟器的重要组成部分，它为导弹导引头提供一定功率、频率的信号，它的正常运行是保证导弹检测设备正常工作的前提，由于通常处于大功率、大电流及高频率工作状态，所以容易出现故障。

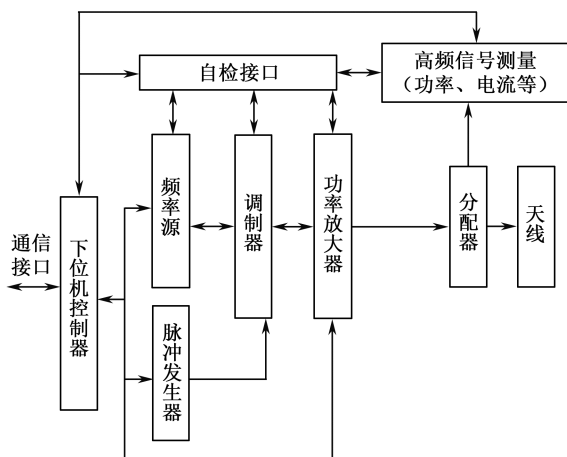


图 8-12 高频模拟器原理框图

导弹检测设备在开机自检和导弹检测过程中，检测结果涉及高频模拟器的状态参数值。另外，通过传感器技术，可以实现对功放组件的状态监测，如输出功率监测、电流监测等。根据其设计原理知，在输入一定的条件下，其中某型功放组件的正常工作输出功率为 $70\sim 74\text{W}$ ，电流为 $2\sim 2.6\text{A}$ ，当超过上述工作范围时，就会出现故障。因此，可将功放组件的输出功率 P 和输出电压 U 作为该功放组件故障预测的特征量。对 3 套功放组件工作状态进行监测，获得的数据如表 8-5 所列。高频模拟器平均每天工作 1.5h。

1. 模型建立

对于同类对象 1 和 2，从第 2 个数据开始的前 7 组数据用于构建预测模型；

对于特定个体 3，从第 1 个数据开始的前 7 组数据用于构建预测模型，后 4 组数据作为预测检验数据。采用基于信息融合和 IMUGM(1,m,w)模型进行预测，使用 Matlab 软件进行仿真计算分析。

表 8-5 某型高频模拟器的状态监测数据

工作时间/h		50	150	250	400	550	700	800	900	950	1000	1050
功率 /W	1	70.28	70.52	70.81	71.16	71.43	71.69	71.86	72.31	72.79	73.33	73.84
	2	70.21	70.33	70.49	70.76	71.12	71.54	71.97	72.35	72.68	73.14	73.56
	3	70.16	70.35	70.69	70.92	71.25	71.58	71.73	72.01	72.24	72.62	73.17
电流/A	1	2.01	2.03	2.05	2.10	2.12	2.19	2.27	2.38	2.46	2.50	2.58
	2	2.06	2.09	2.15	2.17	2.22	2.24	2.28	2.34	2.39	2.46	2.55
	3	2.03	2.04	2.06	2.09	2.14	2.21	2.25	2.32	2.36	2.44	2.47

原始数据序列为

$$X_1^{(0)} = \begin{bmatrix} 70.5200 & 2.0300 \\ 70.8100 & 2.0500 \\ 71.1600 & 2.1000 \\ 71.4300 & 2.1200 \\ 71.6900 & 2.1900 \\ 71.8600 & 2.2700 \\ 72.3100 & 2.3800 \end{bmatrix}, \quad X_2^{(0)} = \begin{bmatrix} 70.3300 & 2.0900 \\ 70.4900 & 2.1500 \\ 70.7600 & 2.1700 \\ 71.1200 & 2.2200 \\ 71.5400 & 2.2400 \\ 71.9700 & 2.2800 \\ 72.3500 & 2.3400 \end{bmatrix}, \quad X_3^{(0)} = \begin{bmatrix} 70.1600 & 2.0300 \\ 70.3500 & 2.0400 \\ 70.6900 & 2.0600 \\ 70.9200 & 2.0900 \\ 71.2500 & 2.1400 \\ 71.5800 & 2.2100 \\ 72.7300 & 2.2500 \end{bmatrix}$$

根据式 (8-21) ~ 式 (8-32)，编程计算得

$$w_{11} = (0.5510, 0.5600, 0.4290)^T, \quad \hat{x}_1^{(0)}(t_8) = (72.1812, 2.3403)$$

$$w_{21} = (0.6036, 0.4511, 0.5379)^T, \quad \hat{x}_2^{(0)}(t_8) = (72.4225, 2.3070)$$

$$w_{31} = (0.4925, 0.5281, 0.4637)^T, \quad \hat{x}_3^{(0)}(t_8) = (71.8630, 2.2963)$$

分别添加 $X_1^{(0)}(t_9)$ 、 $X_2^{(0)}(t_9)$ 、 $X_3^{(0)}(t_8)$ ，并删除 $X_1^{(0)}(t_2)$ 、 $X_2^{(0)}(t_2)$ 、 $X_3^{(0)}(t_1)$ ，计算得

$$w_{12} = (0.6230, 0.5510, 0.5630)^T, \quad \hat{x}_1^{(0)}(t_9) = (72.6522, 2.4905)$$

$$w_{22} = (0.5975, 0.4872, 0.5418)^T, \quad \hat{x}_2^{(0)}(t_9) = (72.4594, 2.4213)$$

$$w_{32} = (0.4396, 0.4641, 0.5169)^T, \quad \hat{x}_3^{(0)}(t_9) = (71.1847, 2.3816)$$

依此类推，直到得到全部数据。

2. 故障预测

根据欧几里得距离 d_{zi} 和隶属度权值 w'_i 的含义，可以编程计算得 $w'_{11} = 0.1713$ ， $w'_{21} = 0.2872$ ， $w'_{31} = 0.5415$ ，进而求得 $\hat{x}^{(0)}(t_8) = (72.0781, 2.3084)$ 。依此类推，最后求得 $w'_{14} = 0.2822$ ， $w'_{24} = 0.1160$ ， $w'_{34} = 0.6018$ ， $\hat{x}^{(0)}(t_{11}) = (73.2598, 2.4851)$ 。

同时，为检验本节预测方法的预测效果，将信息融合和 IMUGM(1,m,w)模型的预测方法与传统 UGM(1,m)模型、改进的单因素预测模型（IMUGM(1,1)模型）和改进的多因素预测模型（IMUGM(1,m,w)模型）的预测结果进行比较分析，各预测方法或模型的预测结果分别如表 8-6 及图 8-13、图 8-14 所示，并分别以预测值（FV）、相对误差（RPE）和平均相对误差（MRE）表示。

表 8-6 不同方法的预测值及相对误差

项目		预测结果												MPE/ %
UGM (1,m)	P	FV	70.1600	70.2183	70.5435	70.8251	71.3397	71.7126	71.8104	71.8842	72.0758	72.3929	72.9835	0.1746
		RPE/%	0.0000	0.1872	0.2072	0.1338	0.1259	0.1852	0.1121	0.1747	0.2273	0.3127	0.2549	
	I	FV	2.0300	2.0551	2.0723	2.0834	2.1087	2.1859	2.2348	2.3342	2.3879	2.4225	2.4481	0.7527
		RPE/%	0.0000	0.7402	0.5971	0.3158	1.4626	1.0905	0.6755	0.6121	1.1822	0.7172	0.8866	
IMU GM (1,1)	P	FV	70.1600	70.3054	70.6238	70.9737	71.3386	71.4820	71.6841	71.9523	72.2934	72.7416	73.2865	0.0913
		RPE/%	0.0000	0.0634	0.0936	0.0757	0.1244	0.1369	0.0640	0.0813	0.0739	0.1674	0.1592	
	I	FV	2.0300	2.0514	2.0687	2.0842	2.1260	2.1829	2.2403	2.2956	2.3467	2.4241	2.4910	0.6079
		RPE/%	0.0000	0.5588	0.4223	0.2775	0.6542	1.2262	0.4311	1.0517	0.5636	0.6516	0.8502	
IMU GM (1,m,w)	P	FV	70.1600	70.2687	70.6431	70.9560	71.1793	71.5025	71.7531	72.0717	72.1944	72.5224	73.0649	0.0906
		RPE/%	0.0000	0.1155	0.0663	0.0508	0.0992	0.1083	0.0313	0.0857	0.0631	0.1344	0.1436	
	I	FV	2.0300	2.0492	2.0675	2.0989	2.1583	2.1956	2.2574	2.2971	2.3734	2.4316	2.4543	0.4986
		RPE/%	0.0000	0.4510	0.3641	0.4258	0.8551	0.6516	0.3289	0.9871	0.5680	0.2172	0.6356	
本节 方法	P	FV	70.1600	70.2826	70.6525	70.9691	71.2627	71.5169	71.7499	72.0781	72.2745	72.6724	73.2598	0.0601
		RPE/%	0.0000	0.0958	0.0530	0.0692	0.0178	0.0882	0.0277	0.0946	0.0477	0.0722	0.1227	
	I	FV	2.0300	2.0479	2.0613	2.0968	2.1285	2.2042	2.2596	2.3084	2.3341	2.4457	2.4851	0.3585
		RPE/%	0.0000	0.3872	0.0631	0.3254	0.5374	0.2624	0.4267	0.5000	0.5974	0.2336	0.6113	

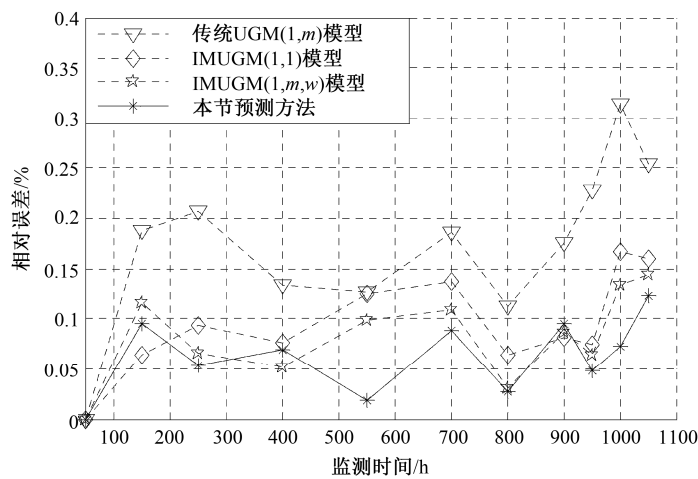


图 8-13 输出功率预测结果对比

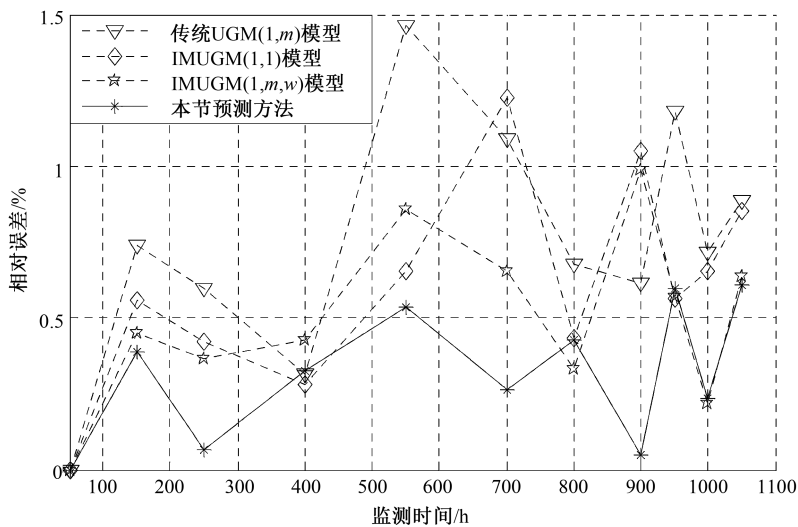


图 8-14 电流预测结果对比

3. 结果分析及应用

从表 8-6、图 8-13 及图 8-14 中可以看出，本节的预测方法拟合效果最好、预测精度最高，其次是 IMUGM(1,m,w)模型，精度最差的是传统 UGM(1,m)模型。IMUGM(1,m,w)模型通过引入加权因子 w 的方式重新构造背景值，并通过新陈代

谢思想对样本数据进行及时更新,还在 IMUGM(1,1)模型的基础上,综合利用了功放组件的多个因素之间的相关性信息;而本节预测方法不但采用了 IMUGM(1,m,w)模型,而且通过信息融合技术充分利用了同类对象的故障信息来提高预测精度。

以第7个工作时间所对应的时间点作为参照,预测下一个备件需求发生的时间。根据上面建立的故障预测模型,可计算出输出功率为 74W、电流为 22.6A 时,工作时间为 1175h,当每天工作 1.5h 时,则有 $(1175-1050)/1.5=83$,即在未来 83 天左右时需要 1 个备件。

8.5 基于故障规律的备件需求预测方法

备件的消耗是由于装备发生故障引发的,备件需求量与装备故障次数之间存在一定的对应关系,通常包括消耗件消耗和可修件消耗两类。装备的故障规律在一定程度上对应着备件的需求规律。掌握了武器装备故障规律,就可以根据故障规律进行备件需求的预测。

8.5.1 备件需求量计算公式

求得装备故障次数预测值,就可以利用它进行备件需求预测。下面分析备件需求量与装备故障次数之间的关系,推导备件需求量计算公式。

1. 故障件流程分析

为了便于问题分析,将故障件分为消耗件和可修件两类。设消耗件和可修件的数量比例分别为 r_x 和 r_k , $r_x + r_k = 1$ 。

备件的消耗源于故障的发生,每更换多少故障件,就要相应消耗多少备件。对于消耗件,其部件一旦使用便成为消耗品;对于可修件,由于损坏程度和修理能力不同,有的部件可修复,有的部件最终不可修复。一般来说,可修复件在修复后将归入周转备件继续使用,相当于消耗的备件得到补充,而不可修复件则要作报废处理,更换的不可修复件部件则属于永久消耗。故障流程图如图 8-15 所示。

如图所示,当某一故障发生时,拆卸下故障件并更换上相应备件,对更换下来的故障件进行检查判断。若为消耗件,则换上的备件成为消耗品,若为可修件,则进行修理。

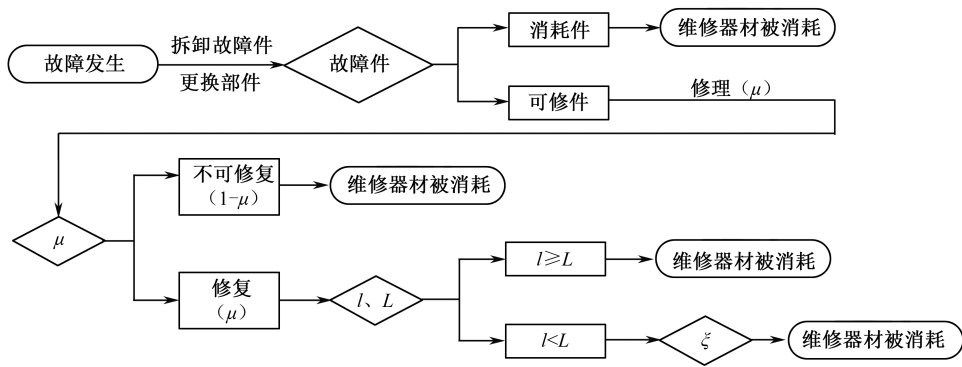


图 8-15 故障流程图

在可修件中，大部分故障件可以修复，小部分则无法修复，修复率 μ 可以作为衡量修复能力大小的指标，不同的部件有不同的修复率值，不同类别也有各自的修复率值。对不可修复件，则换上的备件也成为消耗品；对可修复件，换上的备件是否为永久消耗则要根据修理周期 l 和订货周期 L 的关系来确定。

将二者的关系分为 $l \geq L$ 和 $l < L$ 两种情况，设 $l \geq L$ 和 $l < L$ 两种情况的次数比例分别为 ψ_1 和 ψ_2 ， $\psi_1 + \psi_2 = 1$ ，分别讨论如下：

(1) $l \geq L$ ，即修理周期大于订货周期，则外送修理件在本订货周期内不能返回使用。这样，对于本订货周期内的备件消耗预测而言，该修理件换上的备件则应该认为成为消耗品。

(2) 若 $l < L$ ，即修理周期小于订货周期，则外送修理件在本订货周期内能够返回作为备件使用。这样，对于本订货周期内的备件消耗预测而言，该修理件换上的备件是否成为消耗品要根据不同情况区别对待。

设修理周期与订货周期之比为

$$\xi = \frac{l}{L}$$

修理周期以月为单位，订货周期一般按一年计算，所以，取订货周期 $L = 12$ 个月。

对 $l < L$ 的情况分析如下：

(1) 如果修理周期相对于订货周期而言很短，如 $l \leq 3$ 个月， $\xi \leq 0.25$ ，即一旦可修件发生故障送去修理，在较短的时间内可以返回，作为备件继续使用，这种情况可以近似认为换上的备件没有被消耗。

(2) 如果修理周期相对于订货周期而言很长，如 $l \geq 9$ 个月， $\xi \geq 0.75$ ，即一旦可修件发生故障去修理，在较短的时间内不能够返回，也就无法作为备件继续

使用,因而会给下次使用造成很大不利影响,会严重影响该备件的供应保障,这种情况可以认为换上的备件被消耗掉,成为消耗品。

鉴于以上分析,可以看出,对于 $l < L$,即修理周期小于订货周期的情况,修理周期相对于订货周期的大小程度,即 ξ ($0 < \xi < 1$)的大小,对应着换上的备件成为消耗品可能性大小的程度。 ξ 越大,换上的备件成为消耗品的可能性越大。由此,可以引发这种思想,将 ξ 的大小作为参照系数(表8-7),根据修理周期的具体数值,来确定某件外送修理可修件换上的备件是否成为消耗品和在多大程度上可以认为成为消耗品。

表8-7 参照系数

$l_i/\text{月}$	1	2	3	4	5	6	7	8	9	10	11	12
$\xi_i = l_i/L$	0.08	0.17	0.25	0.33	0.42	0.50	0.58	0.67	0.75	0.83	0.92	1.00

2. 参照系数的确定

不同的修理件有不同的修理周期,由于对同一修理件不同的修理厂家修理能力也不同,因此修理件的参照系数也不同。对于任一类别中的任何一个修理件,按某一标准分为几个大的种类或型类,确定出这些种类或型类的修理件各自的数量比例,再通过对修理件送修和修复返回时间的统计,确定出各自的修理周期,然后计算并查找表8-7中的相应系数,便可以求得该类别备件总的参照系数。具体做法如下。

设 ξ_i 为某一修理周期 l_i ($i=1,2,\dots,12$)对应的参照系数, λ_i 为该修理周期 l_i 对应的某一修理件的数量比例,则该类别备件的总参照系数为

$$\xi = \sum_{i=1}^{12} \xi_i \lambda_i \quad (8-33)$$

3. 五个比率的确

五个比率分别是消耗件比例 r_x 、可修件比例 r_k 、可修件修复率 μ 、 $l \geq L$ 和 $l < L$ 两种情况的次数比例 ψ_1 和 ψ_2 。由于装备保障部门和备件业务部门目前拥有的相关数据资料较少,收集起来难度较大,五个比率值的计算比较困难。对于这五个比率值,可以采用专家调查法,由相关业务部门工作人员根据经验确定。

4. 消耗量计算公式

根据前面分析,对备件消耗量计算公式进行推导,如图8-16所示。

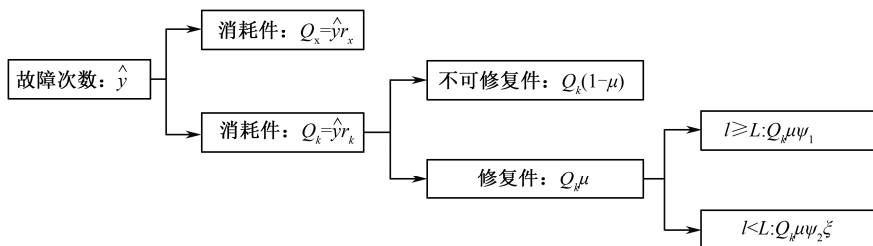


图 8-16 公式推导图

设 \hat{y} 为预测故障次数, Q_x 为消耗件数量, Q_k 为可修件数量, \hat{Q}_x 为消耗件数量预测值, \hat{Q}_k 为可修件数量预测值。则

$$Q_x = \hat{y}r_x, \quad Q_k = \hat{y}r_k$$

由图 8-16 可知: 对消耗件, 消耗数量等于发生故障的数量, 即 $\hat{Q}_x = Q_x$; 对可修件, 消耗数量为不可修复件、 $l \geq L$ 的修复件和 $l < L$ 的修复件三部分数量之和, 即

$$\begin{aligned} \hat{Q}_k &= Q_k(1-\mu) + Q_k\mu\psi_1 + Q_k\mu\psi_2\xi \\ &= Q_k[1 + \mu(\psi_1 + \psi_2\xi - 1)] \end{aligned}$$

令修正系数为

$$\eta = [1 + \mu(\psi_1 + \psi_2\xi - 1)] \quad (8-34)$$

则 $\hat{Q}_k = Q_k\eta$ 。

因为 $0 < \xi < 1$, $\psi_1 + \psi_2 = 1$, 所以 $\psi_1 + \psi_2\xi - 1 < 0$, $\eta < 1$, 因此 $\hat{Q}_k < Q_k$ 。

综上所述, 消耗件和可修件的消耗量计算公式为

$$\hat{Q}_x = \hat{y}r_x \quad (8-35)$$

$$\hat{Q}_k = Q_k\eta = \hat{y}r_k[1 + \mu(\psi_1 + \psi_2\xi - 1)] \quad (8-36)$$

8.5.2 示例分析

已知某型导弹装备陀螺平台 2000—2009 年故障次数数据为 6 次、9 次、7 次、8 次、8 次、10 次、8 次、6 次、8 次、7 次。试确定 2010 年该部件备件的需求量。

1. 故障次数预测

按照前面灰色数列预测步骤, 得灰色预测模型为

$$\begin{cases} \hat{x}^{(1)}(k+1) = -460.5968e^{-0.0186k} + 466.5968 \\ \hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \end{cases}$$

即

$$\hat{x}^{(0)}(k+1) = 8.6473e^{-0.0186k}$$

模型精度为三级，故可用来进行预测。对 2010 年故障次数进行预测：

$$\hat{X}^{(0)} = \hat{x}^{(0)}(11) = 7.1808 \approx 7$$

2. 备件需求量预测

由于陀螺平台属于可修部件，故 $r_k = 1$ 。

根据部件送修数据统计，得 $\mu = 85\%$ ， $l = 6$ 个月。

$l < L$ ，故 $\psi_1 = 0$ ， $\psi_2 = 1$ ， $\xi = 0.50$ 。代入式 (8-34) 和式 (8-36) 计算，可得

$$\eta = 1 + \mu(\psi_1 + \psi_2\xi - 1) = 1 + 85\%(0 + 1 \times 0.50 - 1) = 0.575$$

$$\hat{Q}_k = \hat{y}r_k\eta = 7 \times 1 \times 0.575 = 4.025 \approx 4 \quad (\text{个})$$

第 9 章

基于组合模型的武器装备备件需求预测方法



对于武器装备备件需求的预测，由于历史需求数据在一定程度上存在随机性，单一预测方法都有一定的适用条件或范围，而且运用各种预测模型得到的预测结果精度各不相同，因此，运用组合预测法可以大大提高预测精度。

9.1 组合预测方法概述

近年来，武器装备备件的预测理论和方法有了很大的进展，各种新的理论和方法不断涌现，如果能够把它们有效地组合起来使用，往往能够达到事半功倍的效果。由于组合预测方法和单项预测方法相比能有效地提高预测精度，因此受到国内外预测工作者的重视，组合预测方法已经成为预测领域中的一个重要研究方向。组合预测可根据问题的特征，充分利用各单项预测模型的有用信息，采用两种或两种以上的预测模型，加以适当地有效组合，从而准确、合理地进行预测。组合预测能够综合利用各种预测方法所提供的信息，克服单一模型的局限性，有效地集结更多有用信息。对于武器装备备件需求的预测，由于历史需求数据在一

定程度上存在随机性,单一模型的预测结果的准确性和精度各不相同,而组合预测是可以充分利用各个模型的优点,并选取适当权重进行加权平均的一种预测方法,从而提高了预测的精确度。组合预测模型的关键技术是恰当地确定单个预测模型组合时的权重系数。

针对单一预测方法利用数据和假设条件的局限性,为避免部分可用信息的丢失, Bates 和 Granger 于 1969 年首次提出了组合预测方法的概念,即综合考虑各单项预测方法的优点,将不同的单项预测方法进行组合。20 世纪 70 年代以后,组合预测方法的研究进一步得到了重视,1989 年,国际预测领域的权威学术刊物“*Journal of Forecasting*”出版了组合预测方法专辑。经过对国内外文献的分析,相关研究呈现以下特点:①提出了多种准则下的组合预测模型,对组合预测模型的求解和有效性的实证研究较为深入,但缺乏多种准则框架下组合预测模型有效性的理论研究成果。②对组合预测方法有效性的理论研究已经引起学术界的关注。陈华友等研究了组合预测方法有效度冗余信息的判定等有关内容。③传统的组合预测模型是从改善某种拟合误差角度建立的,同一种单项预测方法在各个时点的权系数是一样的,而实际情况却往往不然。因为同一种单项预测方法在不同的时刻表现并不是完全相同。为了克服这一缺陷,国内外学者相继提出了有序加权平均算子、诱导有序加权平均算子及诱导有序加权调和平均(Induced Ordered Weighted Harmonic Averaging, IOWHA)算子,还与度量预测精度的另一种相关性指标——Theil 不等系数相结合建立组合预测模型,已广泛用于税收预测、用电量预测、需水量预测等领域。④通常情况下,建立组合预测模型都是在模型组合系数上下功夫来降低预测误差,但有时很难达到预期效果。小波变换是近年来兴起的一种新的信号分析处理技术,具有多分辨(也称多尺度)分析的特点,在时域、频域都具有表征信号局部特征的能力,是处理非平稳随机时间序列的一种有效方法。小波变换技术为建立组合预测模型提供了一条新的思路。

9.1.1 一般组合预测模型

对同一预测问题有 n 个预测模型,这 n 个预测模型分别记为 f_1, f_2, \dots, f_n 。再根据对各单一预测模型预测结果的分析,确定出各单一预测模型在组合预测模型中的最优权重 $k_i, i=1, 2, \dots, n$,这样就构成了预测模型:

$$f_t = \sum_{i=1}^n k_i f_{it} \quad (9-1)$$

式中: f_t 为 t 时刻组合预测的预测值; f_{it} 为 t 时刻第 i 种方法的预测值; n 为预测

方法的个数； k_i 为第 i 预测方法的权重且 $\sum_{i=1}^n k_i = 1$ ($0 \leq k_i \leq 1$)。

9.1.2 最优加权系数的确定

最优加权法实质为依据“过去一段时间内组合预测误差平方和最小”原则来求各个预测模型的加权系数。

第 i 种预测模型在第 t 年的预测误差为

$$e_{it} = y_t - f_{it} \quad (i=1, 2, \dots, n) \quad (9-2)$$

式中： y_t 为实际值； t 为统计年 ($t=1, 2, \dots, m$)。

第 i 种预测模型的预测误差向量为

$$\mathbf{F}_i = [e_{i1}, e_{i2}, \dots, e_{in}]^T \quad (9-3)$$

误差矩阵为

$$\mathbf{e} = [\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_n] \quad (9-4)$$

误差信息矩阵为

$$\mathbf{E}_r = \mathbf{e}^T \mathbf{e} = \begin{bmatrix} E_{11} & \cdots & E_{1n} \\ \vdots & \ddots & \vdots \\ E_{n1} & \cdots & E_{nn} \end{bmatrix} \quad (9-5)$$

令 $\mathbf{R}_r = [1, 1, \dots, 1]^T$ 为元素全为 1 的 n 维向量， $\mathbf{W} = [w_1, w_2, \dots, w_n]$ 为 n 种预测模型的权重向量。

则组合预测的误差平方和为

$$S = \sum_{t=1}^m e_t^2 = \sum_{t=1}^m \left(\sum_{i=1}^n w_i e_{it} \right)^2 = \mathbf{W}^T \mathbf{E}_r \mathbf{W} \quad (9-6)$$

利用线性规划求解，即

$$\begin{cases} \min S = \mathbf{W}^T \mathbf{E}_r \mathbf{W} \\ \text{s.t. } \mathbf{R}_r^T \mathbf{W} = 1 \\ \mathbf{W} \geq 0 \end{cases} \quad (9-7)$$

其最优解为

$$\mathbf{W} = \frac{\mathbf{E}_r^{-1} \mathbf{R}_r}{\mathbf{R}_r^T \mathbf{E}_r^{-1} \mathbf{R}_r} \quad (9-8)$$

基于二模型的组合预测，最优加权系数为

$$k_1 = \frac{\sum_{t=1}^m e_{2t}^2 - \sum_{t=1}^m e_{1t} e_{2t}}{\sum_{t=1}^m e_{1t}^2 + \sum_{t=1}^m e_{2t}^2 - 2 \sum_{t=1}^m e_{1t} e_{2t}} \quad (9-9)$$

$$k_2 = \frac{\sum_{t=1}^m e_{1t}^2 - \sum_{t=1}^m e_{1t} e_{2t}}{\sum_{t=1}^m e_{1t}^2 + \sum_{t=1}^m e_{2t}^2 - 2 \sum_{t=1}^m e_{1t} e_{2t}} \quad (9-10)$$

9.1.3 模型的检验

为了进一步检验各模型预测效果的准确性,按照预测效果评价原则和惯例效果评价原则,可以采用拟合误差指标(相对误差和绝对误差)作为评判准则,对预测效果进行全方位的综合评价,通过各预测模型的误差比较,得出最优加权组合预测模型预测精度与单一预测模型精度的关系。

9.2 基于回归分析和灰色模型的备件需求预测方法

目前,对武器装备备件需求预测的方法有时间序列法、灰色 GM(1, 1)方法、回归分析法等,这些方法均采用单项的预测模型进行预测。然而,由于装备故障的复杂性和不确定性使得备件需求受多种因素的共同影响,从而决定了预测效果的波动性,因此,有时单项预测难以准确反映备件需求未来发展的实际情况。

9.2.1 单个预测模型建立

1. 一元线性模型

采用数理统计中的回归分析方法对变量的观测数据进行统计分析,再用最小二乘法求解。建立回归预测模型为

$$y = a + bx \quad (9-11)$$

式中: x 为时间、战备任务量、工作时数等因素; y 为备件消耗量。

2. 灰色 GM(1,1)模型

据灰色系统理论，将备件消耗量这一灰色量累加生成 $x^{(0)}(k)$ 序列，建立 GM(1,1)模型为

$$\begin{cases} \hat{x}^{(0)}(1) = \hat{x}^{(1)}(1) \\ \hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = \left(\hat{x}^{(0)}(1) - \frac{u}{a} \right) (1 - e^a) e^{-a(k-1)} \quad (k=2,3,\dots,n) \end{cases} \quad (9-12)$$

9.2.2 最优组合预测模型

设 f_1 为一元线性回归预测值； f_2 为灰色模型预测值； k_1 和 k_2 分别是一元线性回归模型和灰色模型最优加权系数，且 $k_1 + k_2 = 1$ ，则最优组合预测模型为

$$f_t = k_1 f_1 + k_2 f_2 \quad (9-13)$$

9.2.3 示例分析

某型自动化测试设备 2001—2009 年的累计通电时间及印制板的消耗情况，见表 9-1。预计至 2010 年装备累计通电时间达到 4855h，试确定至 2010 年该部件备件需求量。

表 9-1 某型自动化测试设备印制板备件消耗数据统计

年份	2001	2002	2003	2004	2005	2006	2007	2008	2009
装备通电时间/h	1077	1228	1581	1883	2286	2387	2739	3294	3389
备件消耗量/个	17	19	22	34	38	42	51	61	71

基于最优组合预测模型的建模思想，选取 2001—2008 年实际累计消耗量作为原始数据，分别建立一元线性回归模型、GM(1,1)模型和最优加权组合预测模型，用 2009 年数据作为检验数据，预测结果见表 9-2。

1. 一元线性回归模型

根据式 (9-11) 可得

$$y = -6.38 + 0.01985x$$

式中： $a = -6.38, b = 1.985$ 。

标准离差 $s = 0.022$ ， $\frac{s}{\bar{y}} = 6.14\% < 15\%$ （ \bar{y} 为估计值 \hat{y}_i 与实际值 y_i 的平均误差）。相关系数 $r = 0.994$ ，拟定显著性水平 $\alpha = 0.05$ ，查相关系数检验表，取 $\nu = n - 2 = 8 - 2 = 6$ ，得到 $r_n = 0.7067$ ， $|r| \geq r_n$ 显著相关，检验通过。

表 9-2 各模型拟合情况

年份	实际值 /个	一元线性回归			GM(1,1)			最优加权组合		
		拟合值 /个	相对误差 /%	绝对 误差	拟合值 /个	相对误差 /%	绝对 误差	拟合值 /个	相对误差 /%	绝对 误差
2001	17	15	-11.76	-0.02	21	23.53	0.04	17	0	0
2002	19	18	-5.26	-0.01	23	21.05	0.04	20	5.26	0.01
2003	22	25	13.64	0.03	26	16.18	0.04	25	13.64	0.03
2004	34	31	-8.82	-0.03	34	0	0	32	-5.89	-0.02
2005	38	39	2.56	0.01	39	2.56	0.01	39	2.63	0.01
2006	42	41	-2.44	-0.01	44	4.76	0.02	42	0	0
2007	51	48	-6.25	-0.03	50	-1.96	-0.01	49	-4.56	-0.02
2008	61	59	3.39	-0.02	57	-6.56	-0.04	58	-4.38	-0.03

2. GM(1,1)模型

建立灰色 GM(1,1)模型得到： $\hat{a} = -12.97, \hat{u} = 19.94$ 。

根据式 (9-12) 的 GM(1,1)方程：

$$\hat{x}^{(0)}(k+1) = (1 - e^{-12.97})(17 + 1.5374)e^{12.97k}$$

小误差概率 $P = |\varepsilon(k) - \bar{\varepsilon}| = 1 > 0.95$ ，模型精度为一级。 $c = 0.10 < 0.35$ 。

3. 最优加权组合预测模型

由表 9-2 中计算的各单一模型误差结果，利用式 (9-9)、式 (9-10) 计算最优加权系数： $k_1 = 0.664, k_2 = 0.336$ 。

根据式 (9-13) 可得预测模型为

$$f_t = 0.664f_1 + 0.336f_2$$

图 9-1 中各模型的备件消耗量拟合情况表明，一元线性回归、GM(1,1)拟合的误差较大，最优加权预测法的效果最好，其结果拟合出的曲线与实际值曲线几乎重合。

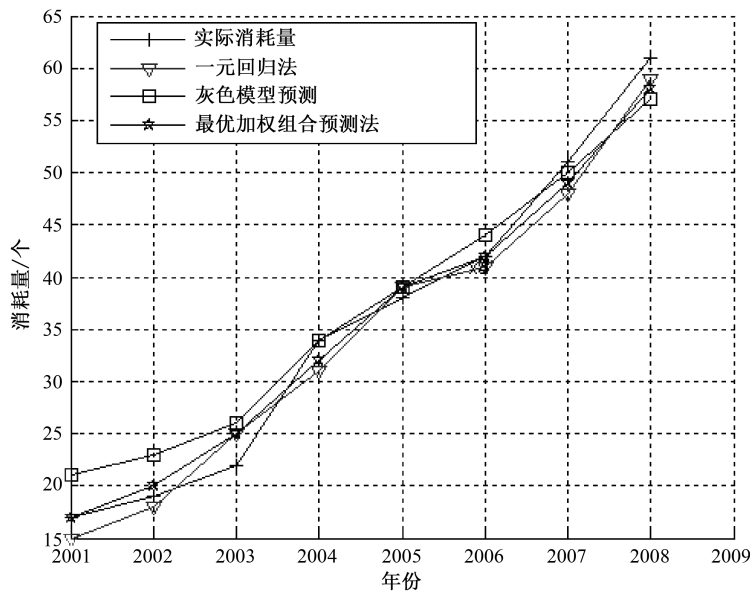


图 9-1 各模型的预测精度

4. 模型评价及预测

为检验最优加权组合模型的预测效果，对 2009 年备件累计消耗量进行了预测，见表 9-3。由表 9-3 可知，最优加权组合预测模型的结果与实际值完全吻合，误差为零，精度高，与一元线性回归模型、GM(1,1)模型比较提高了预测的准确性，改善了预测效果。

利用建立的最优组合预测模型对 2010 年备件需求量进行预测，结果为 90 个。

表 9-3 2009 年备件需求预测

预测模型	实际值/个	预测值/个	相对误差/%	绝对误差
一元线性回归	71	70	-1.41	-0.01
GM(1,1)		72	1.41	0.01
最优加权组合		71	0	0

9.3 基于小波变换的备件需求预测方法

很多武器装备备件需求受其自身可靠性、使用强度、环境条件、人为操作等多种复杂因素的影响,因此,其需求规律大都呈现非线性、非平稳特征。小波变换是近年来兴起的一种新的信号分析处理技术,具有多分辨分析的特点,在时域、频域都具有表征信号局部特征的能力,是处理非平稳随机时间序列的一种有效方法。目前,小波变换广泛应用于数值分析、信号处理、图像处理、计算机视觉等领域,但用于备件需求预测方面的研究还几乎没有。为了解决备件需求非平稳时间序列的预测问题,本书将小波变换技术引入备件需求预测领域,并构建了基于小波变换和改进 GM(1,1)-ARMA 模型的备件需求预测模型(简称 WGM-ARMA 模型)。

9.3.1 小波变换理论

1. 小波变换基本理论概述

小波变换通过有限长的平移小波函数可获得信号的时间信息,通过缩放小波的宽度(尺度)可获得信号的频率特性。小波变换可分为离散小波变换和连续小波变换两大类。由于连续小波变换的计算量非常大,因此在实际工程应用中经常采用离散小波变换。

设函数 $\psi(t) \in L^2(\mathbf{R})$ ($L^2(\mathbf{R})$ 表示平方可积的实数空间), 其傅里叶变换 $\hat{\psi}(w)$ 满足容许条件

$$C_{\psi} = \int_{\mathbf{R}} \frac{|\hat{\psi}(w)|}{|w|} dw < +\infty \quad (9-14)$$

则称 $\psi(t)$ 为母小波。将母小波经过伸缩和平移后, 就得到一个小波序列

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (a, b \in \mathbf{R}, a \neq 0)$$

式中: a 为伸缩因子; b 为平移因子。

设 $f(t) \in L^2(\mathbf{R})$, $\psi(t)$ 为基本小波, 则函数 $f(t)$ 的连续小波变换为

$$W_f(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t) \bar{\psi}\left(\frac{t-b}{a}\right) dt \quad (9-15)$$

式中: $\bar{\psi}(t)$ 为 $\psi(t)$ 的复共轭函数。

通常, 把连续小波变换中参数 a 和 b 的离散化公式分别取作 $a = a_0^m$ ($a_0 > 1, m \in \mathbf{Z}$), $b = na_0^m b_0$ ($b_0 > 0$)。这时, 对应的离散小波函数为

$$\psi_{a,b}(t) = a_0^{-m/2} \psi\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) = a_0^{-m/2} \psi(a_0^{-m} t - nb_0)$$

及其对应的离散小波变换为

$$W_f(a_0^m, na_0^m b_0) = \int_{\mathbf{R}} f(t) a_0^{-m/2} \bar{\psi}(a_0^{-m} t - nb_0) dt = \langle f, \psi_{m,n} \rangle \in \mathbf{Z}^2 \quad (9-16)$$

2. 小波分解与重构技术研究

小波分解与重构有很多算法, 常用快速变换算法: Mallat 算法 (塔式算法) 和 A. Trous 算法等。本书根据备件需求时间序列特点, 采用 Mallat 算法。Mallat 算法从空间概念方面形象地说明了多分辨分析特性, 并给出了信号分解不同频率通道的算法及重构算法。

1) 小波基的选择

在进行小波分解时, 小波函数的选择通常遵循以下原则:

(1) 满足一定区间的时域紧支撑性。适当的时域支撑长度可以保证小波函数在频域具有合适的通频带宽, 使小波分解时具有较好的高频和低频分离能力。支撑宽度越小, 小波基的局部化能力越强, 小波变换的计算复杂度越低, 而支撑越大, 正则性越好。

(2) 较高的消失矩和较好的正则性。较高的消失矩和较好的正则性使小波函数具有适当的低频平滑能力。

(3) 正交性或近似正交性。正交性可使通过小波分解得到的各分量保持一定的独立性。

(4) 对称性。对称的小波基具有线性相位, 对称滤波器组使边缘附近对称的量化误差不敏感, 并可避免信号分解和重构时的相位失真。

常用小波函数的性质如表 9-4 所列。

实际预测过程中, 可根据不同的问题选择不同的小波母函数。由于进行备件需求预测时, 对象是具有趋势性和波动性的时间序列, 因此要求小波函数具有以下特性: ①良好的紧支撑性; ②较高的消失矩; ③良好的正则性。显然, Symlet 小波具有以上特性, 且实践表明 Symlet 小波能够较好地分析时间序列问题。

表 9-4 常用小波函数的性质

小波函数	正交性	紧支撑性	支撑长度	正则性	对称性	消失矩
Haar	有	有	1	$\alpha = 1$	对称	1
Daubechies	有	有	$2N-1$	随 N 增大	近似对称	N
Symlet	有	有	$2N-1$	随 N 增大	近似对称	$2N$
Meyer	有	无	有限长度	$\alpha \in [0, \infty)$	对称	—
Coiflet	有	有	$6N-1$	随 N 增大	近似对称	$2N$
BiorSplines	无	有	$2N$ 或 2^N	$\alpha \geq N$	不对称	$N-1$

2) 小波分解层数的确定

在小波多分辨分析中,分解级数的选取是比较关键的。若所选尺度较少,则不能有效地分析备件需求时间序列特性;所选尺度过多,则增加了预测的复杂度和难度。本书通过综合考虑时间序列的均方根差变化量、互相关系数、信噪比及平滑度四个评价指标来确定最佳分解级数。

小波 M 级分解与重构的时间序列均方根差为

$$\text{RMS}(M) = \sqrt{\frac{\sum_{i=1}^n (x(i) - \hat{x}_M(i))^2}{n}} \quad (9-17)$$

式中: $\hat{x}_M(i)$ 为 M 级分解时间序列; $x(i)$ 为原始时间序列; n 为时间序列长度。

相邻两级的时间序列均方根差变化量为

$$v(M) = \text{RMS}(M+1) - \text{RMS}(M) \quad (9-18)$$

$v(M)$ 反映了相邻两级分解与重构数据时间序列的相似程度,评价准则为均方根差变化量越接近 0 越好。

互相关系数为

$$\rho(M) = \text{cov}(\hat{x}_M(i), x(i)) / (\sigma_x \cdot \sigma_{\hat{x}_M}) \quad (9-19)$$

式中: σ_x 、 $\sigma_{\hat{x}_M}$ 分别为原始时间序列和 M 级分解重构时间序列的方差。评价准则为互相关系数越接近 1 越好。

信噪比为

$$\text{SNR}(M) = 10 \lg \left(x(i)^2 / (x(i) - \hat{x}_M(i))^2 \right) \quad (9-20)$$

信噪比为数据时间序列的能量与噪声能量之比,评价准则为信噪比越大越好。

平滑度为

$$r(M) = (\hat{x}_M(i+1) - \hat{x}_M(i))^2 / (x(i+1) - x(i))^2 \quad (9-21)$$

该项指标可以反映重构后时间序列的平滑程度,评价准则为平滑度越小,异常数据处理效果越好。

用上述四个评价指标中的某几个确定最佳级数都有片面性,综合运用四个指标可有效体现数据的整体特性。把时间序列均方根差变化量、互相关系数、信噪比、平滑度分别进行归一化处理,公式为

$$P_v(M) = [v_{\max} - v(M)] / (v_{\max} - v_{\min}) \quad (9-22)$$

$$P_\rho(M) = [\rho(M) - \rho_{\min}] / (\rho_{\max} - \rho_{\min}) \quad (9-23)$$

$$P_{\text{SNR}}(M) = [\text{SNR}(M) - \text{SNR}_{\min}] / (\text{SNR}_{\max} - \text{SNR}_{\min}) \quad (9-24)$$

$$P_r(M) = [r_{\max} - r(M)] / (r_{\max} - r_{\min}) \quad (9-25)$$

将各评价指标归一化值相加得到总体评价指标为

$$H(M) = P_v(M) + P_\rho(M) + P_{\text{SNR}}(M) + P_r(M) \quad (9-26)$$

归一化后的分指标值越大,对应各分指标效果越好,所以总体评价指标 $H(M)$ 越大,则去噪效果越好, $H(M)$ 最大值所对应的级数 M 即最佳级数。可以看出,本书提出的方法既包含了更多的评价指标,又能定量确定最佳级数,相对已有方法更加合理和严密。

9.3.2 GM(1,1)-ARMA 模型的改进

1. 改进的 GM(1,1)预测模型

灰色预测模型作为灰色系统理论的基本模型之一,是建立系统运行趋势模型的有效方法,适用于动态预测,且只需知道少量已知信息就可建立预测模型。它是以灰色模块为基础,以微分拟合法为工具而建成的模型。目前使用最广泛的灰色预测模型为单变量一阶灰色预测模型 GM(1,1),它通过对原始序列进行累加生成操作后建立一阶线性微分方程模型,然后利用累减生成操作还原为原始序列的预测值。

1) 初始值的改进

传统的 GM(1,1)模型在求解灰微分方程时,初始条件为 $\hat{x}^{(1)}(1) = x^{(1)}(1) = x^{(0)}(1)$,即拟合曲线必然通过第一个点 $(1, x^{(0)}(1))$ 。该方法缺少理论依据,在后面的预测中会降低模型的模拟精度和预测精度。本节运用灰色自适应模型的思想是对灰色模型中的初始值进行修正。

设能达到最优化的模型的初始值为 $x^{(1)}(1) + c$, c 为一个未知优化量,则白

化微分方程的时间响应函数变为

$$\hat{x}^{(1)}(k) = \left[x^{(1)}(1) + c - \frac{b}{a} \right] e^{-a(k-1)} + \frac{b}{a} \quad (k=1, 2, \dots) \quad (9-27)$$

对于优化量 c ，可通过使生成序列新预测值的误差在最小二乘意义下最小来求解，即

$$U(c) = \sum_{k=1}^n [\hat{x}^{(1)}(k) - x^{(1)}(k)]^2 \quad (9-28)$$

为了使 $U(c)$ 最小，令 $\frac{\partial U}{\partial c} = 0$ ，求偏导得

$$c = \left\{ \gamma - \xi \left[x^{(0)}(1) - \frac{b}{a} \right] - \frac{b}{a} \eta \right\} / \xi \quad (9-29)$$

式中

$$\xi = (1 - e^{-2na}) / (1 - e^{-2}), \quad \eta = (1 - e^{-na}) / (1 - e^{-a}), \quad \gamma = \sum_{k=1}^n x^{(1)}(k) e^{-a(k-1)}$$

2) 背景值的改进

由 GM(1,1)模型的建模过程可知，模型的拟合和预测精度主要取决于系数 a 和 b ，而 a 和 b 的求解依赖于背景值 $z^{(1)}(k)$ 的构造形式，因此，背景值 $z^{(1)}(k)$ 的值是直接影响 GM(1,1)模型精度和适应性的主要因素。

由 GM(1,1)模型建模过程可知， $x^{(1)}(t)$ 在区间 $[k, k+1]$ 上的背景值 $z^{(1)}(k+1)$ 是 $x^{(1)}(k)$ 与 $x^{(1)}(k+1)$ 的平均值，由背景值构造示意图 9-2 可知， $z^{(1)}(k+1)$ 为梯形 $ABCD$ 的面积。实际上， $x^{(1)}(t)$ 与横坐标轴所围曲面的面积才是背景值 $z^{(1)}(k+1)$ 。由于 GM(1,1)模型拟合曲线是指数曲线，所以在 $[k, k+1]$ 上实际曲线 $x^{(1)}(t)$ 对应的面积总是小于梯形面积，面积差为 ΔS 。当 $x^{(1)}(t)$ 为低增长指数曲线时， ΔS 比较小，用梯形面积值作为 $z^{(1)}(k+1)$ 时，模型偏差较小。但当 $x^{(1)}(t)$ 为高增长指数曲线时，用梯形面积值作为 $z^{(1)}(k+1)$ 值时，模型偏差较大。

由于原始数据是离散的数据序列，实际曲线 $x^{(1)}(t)$ 也未知，曲面的实际面积就无法计算。为此，可将区间 $[k, k+1]$ 等分为 n 个小区间，用这 n 个小区间的面积和表示实际面积。在 n 较小时， n 个小区间面积和小于实际面积；在 n 较大时， n 个小区间面积和大于实际面积。因此，理论上存在一个 n 值，使 n 个小区间的面积和等于实际面积。如果将该 n 值对应的小区间面积之和作为 $z^{(1)}(k+1)$ ，GM(1,1)模型偏差最小，其拟合和预测精度将更高。

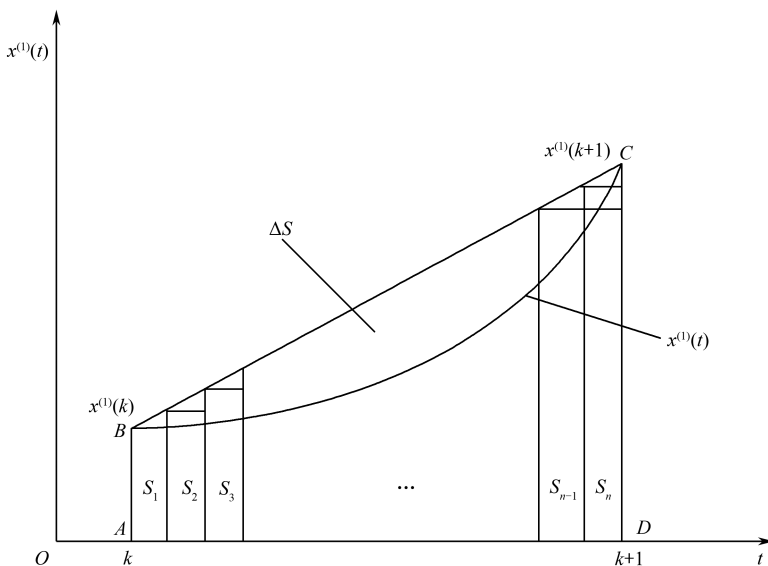


图 9-2 背景值构造

上述改造背景值的方法虽然取得了较好的效果,但其计算过程复杂。 $GM(1,1)$ 模型背景值 $z^{(1)}(k)$ 通常是按照数值积分的梯形公式构造,即将式 (6-2) 在区间 $[k, k+1]$ 上积分,有

$$x^{(1)}(k+1) - x^{(1)}(k) + a \int_k^{k+1} x^{(1)}(t) dt = b(k+1-k)$$

即

$$x^{(0)}(k) + a \int_k^{k+1} x^{(1)}(t) dt = b(k=1, 2, \dots, n-1) \quad (9-30)$$

将式(9-30)和式(6-3)比较可知,用一阶线性微分式(6-2)的解来逼近 $x^{(1)}(k)$,其误差来源于用 $\frac{1}{2}(x^{(1)}(k-1) + x^{(1)}(k))$ 代替 $\int_k^{k+1} x^{(1)}(t) dt$ 。为消除此误差,记

$$z^{(1)}(k) = \int_k^{k+1} x^{(1)}(t) dt \quad (9-31)$$

由于方程式 (6-2) 的解为指数形式,所以可将 $x^{(1)}(t)$ 表示为指数曲线,即

$$x^{(1)}(t) = Ae^{Bt} \quad (9-32)$$

并假设该曲线过 $x^{(1)}(k+1)$ 和 $x^{(1)}(k)$ 两点,因此有

$$x^{(1)}(k) = Ae^{Bk} \quad (9-33)$$

$$x^{(1)}(k+1) = Ae^{B(k+1)} \quad (9-34)$$

解式 (9-33) 和式 (9-34) 得

$$A = \frac{x^{(1)}(k)}{e^{Bk}} = \frac{[x^{(1)}(k)]^{k+1}}{[x^{(1)}(k+1)]^k} = Ae^{B(k+1)}, B = \ln \left[\frac{x^{(1)}(k+1)}{x^{(1)}(k)} \right] = \ln x^{(1)}(k+1) - \ln x^{(1)}(k)$$

因此, 背景值为

$$z^{(1)}(k+1) = \int_k^{k+1} x^{(1)}(t) dt = \int_k^{k+1} Ae^{Bt} dt = \frac{A}{B} [e^{B(k+1)} - e^{Bk}] = \frac{x^{(1)}(k+1) - x^{(1)}(k)}{\ln x^{(1)}(k+1) - \ln x^{(1)}(k)} \quad (9-35)$$

但是, 由式 (9-27) 和式 (9-32) 可知, 式 (9-32) 在利用指数曲线表示 $x^{(1)}(t)$ 的过程中省略了 b/a 项。严格来讲, b/a 只有在 $|a| \geq b$ 时, 才可以省略。通常情况下, 在 $x^{(1)}(t)$ 的表达式中省略常数项 b/a , 将影响背景值对 $\int_k^{k+1} x^{(1)}(t) dt$ 的逼近效果。因此, 在上述背景值改造方法中, 保留 $x^{(1)}(t)$ 的指数表达式中 b/a 项, 以提高 GM(1,1) 模型的适应范围和拟合、预测精度, 则有

$$x^{(1)}(t) = Ae^{Bt} + C \quad (9-36)$$

式中: A 、 B 和 C 为待定常数。

由式 (9-31) 和式 (9-36) 可得

$$z^{(1)}(k) = \frac{x^{(1)}(k) - x^{(1)}(k-1)}{B} + C \quad (9-37)$$

由式 (9-36) 可得

$$\frac{x^{(1)}(k) - C}{x^{(1)}(k-1) - C} = \frac{Ae^{Bk}}{Ae^{B(k-1)}} = e^B$$

则有

$$B = \ln [x^{(1)}(k) - C] - \ln [x^{(1)}(k-1) - C] \quad (9-38)$$

将式 (9-36) 进行累减还原, 可得

$$x^{(0)}(k) = Ae^{Bk} + C - [Ae^{B(k-B)} + C] = Ae^{Bk} \cdot (1 - e^B) \quad (9-39)$$

则累加序列为

$$x^{(1)}(k) = A(1 - e^{-B}) \cdot \sum_{i=1}^k e^{Bi} = Ae^{Bk} - A \quad (9-40)$$

比较式 (9-36) 和式 (9-40) 可得, $C = -A$ 。

将 $k=1$ 和 $C=-A$ 代入式 (9-36), 可得

$$A = \frac{x^{(1)}(1)}{e^B - 1} \quad (9-41)$$

由式 (9-38) 和式 (9-41) 可得

$$B = \ln \left[x^{(1)}(k) - x^{(1)}(1) \right] - \ln x^{(1)}(k-1), \quad C = -A = \frac{x^{(1)}(1) \cdot x^{(1)}(k-1)}{x^{(1)}(1) - x^{(0)}(k)}$$

所以, 新背景值构造公式为

$$z^{(1)}(k) = \frac{x^{(0)}(k)}{\ln \left[x^{(1)}(k) - x^{(1)}(1) \right] - \ln x^{(1)}(k-1)} + \frac{x^{(1)}(1) \cdot x^{(1)}(k-1)}{x^{(1)}(1) - x^{(0)}(k)} \quad (9-42)$$

用式 (9-42) 代替式 (6-4) 的背景值构造公式, 建模过程同 GM(1,1) 模型。

2. 改进动态自适应隔代映射遗传算法 (IDA-IP-GA)

隔代映射遗传算法 (Intergeneration Projection Genetic Algorithm, IP-GA) 最初是在小种群遗传算法的基础上提出的。IP-GA 除继承了小种群遗传算法能避免早熟收敛的优点外, 又比传统遗传算法具有更快的全局收敛速度, 还在进化过程中加入了一个简单的启发式操作算子——隔代映射算子。本书在具有较强局部搜索能力的 IP-GA 基础上, 通过引入自适应策略, 进行种群个体编码, 自行调整个体的适应度, 进而根据种群适应度调整交叉和变异概率, 有效避免了早熟问题, 提高了收敛速度与解的质量, 本书中称为改进动态自适应隔代映射遗传算法 (Improved Dynamic Adaptive IP-GA, IDA-IP-GA), 其流程如图 9-3 所示。

1) 遗传算法的自适应编码

为了推动种群的进一步进化, 可改变个体编码, 假设 GA 所产生的当前种群 $X(t_1)$ 已大体稳定到了齐次种群 $[\bar{X}]$ (或它的均值稳定到 \bar{X}), 则选取小参数 $\delta > 0$, 将此后的解码变换限定到 $e^{-1}: H_L \rightarrow B(\bar{X}, \delta)$ 。这里, $B(\bar{X}, \delta) = \{X: \|X - \bar{x}\|_{\infty} \leq \delta\}$, 是以 X 为中心、以 2δ 为边长的超正方体。比如, 在使用二进制编码时, 要求在 t_1 代之后使用解码变换:

$$x_i = \left[e^{-1}(A) \right]_i = \left[e^{-1}(\bar{X}) \right]_i + \left(\sum_{j=1}^{m_i} b_{ij} 2^{j-1} \right) \frac{2\delta}{2^{m_i} - 1} \quad (9-43)$$

在 GA 中, 一旦确定了编码方案, 就确定了最优值点在编码空间的位置。若采用动态编码, 则 GA 有可能采用具有较快收敛速度的编码方案, 减少编码方案对 GA 收敛性的影响。

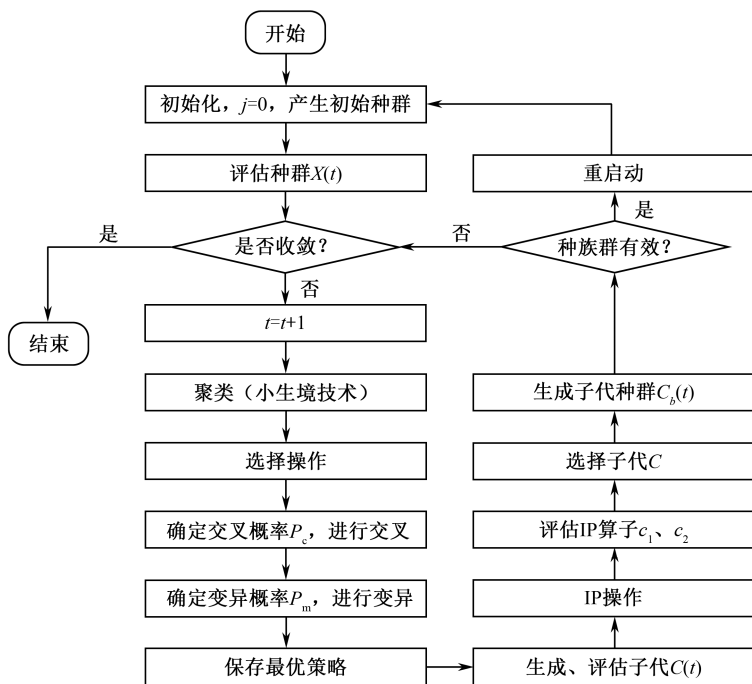


图 9-3 IDA-IP-GA 算法流程

2) 适应度选择策略与自适应算法

适时利用 GA 搜索过程中获得的信息调整个体的适应度, 可以引导 GA 搜索朝着更接近全局最优解方向演化。假设已经知道 GA 到当前代进化中所发现的最佳个体是 $X^*(t)$, 则在下一代的进化中采用适应度度量则可起到压制不优于已发现的 $X^*(t)$ 的个体重复复制的效果; 取时变的适应度度量 $J_t(X) = \exp(\alpha(t)f(X))$, 则可起到在 GA 前期压制特别占优个体过度繁殖, 保持种群多样性, 而在后期强化局部搜索保证收敛性的效果。当处理约束优化问题 $\max_{X \in \Omega} f(X)$ 时, 假设 $\Omega = \{X \in \mathbf{R}^n | g_i(X) = 0, i = 1, 2, \dots, K\}$, 常要用罚函数方法将约束问题化归到一个无约束问题求解, 即通过求解

$$\max f_{\alpha}(X) = f(X) + \sum_{i=1}^K \alpha_i g_i^2(X)$$

来达到求解 $\max_{X \in \Omega} f(X)$ 的目的, 这里 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$ 是一组非常大的正参数(称为罚因子)。

3) 选择算子的自适应策略

作为动态自适应选择的例子,考虑对二进制编码遗传算法的父子混合选择算子进行动态修正。设 $X(t)$ 是父代种群, $X'(t)$ 是子代种群 (它经由 $X(t)$, 通常 GA 一步生成)。令 $\lambda(X(t))$ 表示向量 $\sum_{i=1}^N X_i(t)$ 取非零值和不等 N 的个数 (称为 $X(t)$ 的多样性), $J_{\text{op}}(X) = \max J(X_j(t))$ 表示种群 X 的最佳适应度。

令

$$C_t = \min \left\{ 1, \left[\frac{\lambda(X'(t))}{\lambda(X(t))} \right]^{\frac{\alpha}{t}} \left[\frac{J_{\text{op}}(X'(t))}{J_{\text{op}}(X(t))} \right]^t \right\}, \quad Y(t) = \begin{cases} X'(t) & \text{random}(0,1) < C_t \\ X(t) & \text{random}(0,1) \geq C_t \end{cases}$$

则动态父子混合选择 $S\mathcal{E}_s(t): H_L^N \times H_L^M \rightarrow H_L^M$ 定义为 $S\mathcal{E}_s(t)(X(t) \vee X'(t)) = S(Y(t))$ 。这里, S 是任一确定的比例型选择算子。如此定义的动态父子混合选择能有效地克服 GA 假收敛现象,相应 GA 能依概率收敛到问题的全局最优解。

4) 交叉与变异算子的自适应策略

在标准的遗传算法中,交叉概率 p_c 和变异概率 p_m 在整个进化进程中保持不变,是导致算法性能下降的重要原因。朱金钧等给出了有关自适应交叉概率的基本思想:为防止优良基因结构遭到破坏,须使适应度大的解有较小的 p_c , 使适应度小的解有较大的 p_c , 即 p_c 与 $F_{\max} - \bar{F}$ 成正比。但这一思想未体现出与进化代数的关系。本书在其基础上提出了以下自适应交叉概率:

$$P_{c(t)} = \sqrt{\frac{1}{2}(FA^2 + FC^2)} \cdot \left(1 - \frac{t}{T}\right) \quad (9-44)$$

式中: t 为遗传代数; T 为终止代数。

FA 和 FC 分别为

$$FA = \frac{F_{\max} - \bar{F}}{F_{\max} - F_{\min}}; \quad FC = \frac{F_{\max} - F(x)}{F_{\max} - F_{\min}}$$

其中: F_{\max} 为最大适应度值; F_{\min} 为最小适应度值, \bar{F} 为平均适应度值, $F(x)$ 取 $F(x_i)$ 和 $F(x_j)$ 之间的较大值, x_i 和 x_j 为两个待交叉的个体。

这里所给出的这种交叉概率能够保证在迭代初期交叉概率较大,从而加快进化的速度,避免 GA 陷入迟钝状态;同时能够保证在迭代后期交叉概率较低,并逐步减小至一常量从而保证平滑收敛。

同样,利用人们获得的变异概率的启发式知识和经验,本书提出了如下的自

适应变异概率:

$$P_{m(t)} = \frac{FM}{1 + FA} \left(1 - \frac{D}{D_{\max}} \right) \cdot t^{-\frac{1}{L}} \quad (9-45)$$

式中: t 为遗传进化代数; L 为染色体的长度; D 为父代之间的欧几里得距离; FM 为 $FM = \frac{F_{\max} - F(x_m)}{F_{\max} - F_{\min}}$, 其中 $F(x_m)$ 为变异个体 x_m 的适应度值。

这里所给出的这种自适应变异概率能够保证在迭代初期, 个体性能较差时, 变异概率稍大, 以造成足够的扰动、扩大解空间。随着迭代次数的增加, 该变异概率逐步减小至一个常量从而保证平滑收敛。

3. 基于 IDA-IP-GA 和阻尼最小二乘法优化的 ARMA

自回归移动平均模型是分析平稳随机信号的一种成熟而有效的方法, 能处理周期性变化的序列, 具有统计上的完善性和牢固的理论基础, 且建模时所需数据量少, 预测精度较高, 广泛用于时间序列分析、信号处理和状态估计等领域。ARMA(p, q) 模型是有限参数线性模型, 形式如下:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (9-46)$$

式中: x_t 为时间数据序列; p 、 q 为模型阶数; $\phi_1, \phi_2, \dots, \phi_p$ 为自回归系数; $\theta_1, \theta_2, \dots, \theta_q$ 为移动平均系数; $\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$ 为残差序列 (白噪声序列)。

ARMA 模型参数的估计过程是复杂的非线性回归过程, 导致 ARMA 模型阶数的确定与参数的估计成为模型运用的一个难点。

1) 模型阶数的确定

由于 ARMA 模型的定阶过程存在一定的随机性和不确定性, 为此, 需要综合考虑模型的各种制约因素。本书结合赤池信息 (Akaike Info Criterion, AIC) 准则, 采用 IDA-IP-GA 对 ARMA 模型阶数进行优化, 以期提高 ARMA 模型的预测精度。具体优化过程如下:

首先设定迭代次数、初始种群及数据长度等基本信息。综合考虑实时性、有效性等要求, 设定迭代次数为 100 代, 初始种群数量为 40 组, 每组数据长度为 6 位二进制数, 即前 3 位代表自回归阶数 p , 后 3 位代表移动平均阶数 q 。将随机生成的初始种群进行解码, 即可得到 40 组自回归阶数和移动平均阶数的十进制数组合:

$$(b_0 b_1 \cdots b_n)_2 = \left(\sum_{i=0}^n b_i 2^i \right)_{10}$$

进而确定 ARMA 模型的具体阶数, 并通过 `armax` 函数进行参数估计, 可得此次迭代循环下每组阶数组合的预测结果。

通过 IDA-IP-GA 计算 ARMA 模型阶数, 如果预测值与实际值相差越小, 说明预测参数越接近参数真值。通常衡量预测值与实际值差别的变量采用平均相对变动值 (Average Relative Variance, ARV), 其定义为

$$V_{AR} = \frac{\sum_{i=1}^N [\hat{x}_i - x_i]^2}{\sum_{i=1}^N [\hat{x}_i - \bar{x}]^2} \quad (9-47)$$

式中: N 表示比较数据个数; x_i 为实际数据值; \bar{x} 为实际数据平均值; \hat{x}_i 为预测值。

显然, 平均相对变动值 V_{AR} 越小, 表明选定参数越好。当 $V_{AR} = 0$ 时, 表示达到了理想预测效果; 当 $V_{AR} = 1$ 时, 表明模型仅达到平均值的预测效果。根据适应度值越大越利于遗传到下一代的特性, 可以构造如下适应度函数:

$$f(x) = \frac{1}{V_{AR}} \quad (9-48)$$

利用适应度函数得出适应度函数值最大的个体, 即种群中最好的个体, 将其无条件地选择到下一代新种群中, 然后对父代种群进行复制、交叉和变异等遗传算子运算, 从而繁殖出下一代新种群中的其他 $N-1$ 个基因串。采用联赛法作为复制方法。交叉和变异是产生新个体的遗传算子: 交叉率太大, 将使高适应度的基因串结构很快被破坏掉, 交叉率太小, 则使搜索停止不前, 一般取 $0.5 \sim 0.9$ (本书交叉率取为 0.6); 变异率太大, 将使遗传算法变为随机搜索, 太小则不会产生新个体, 一般取为 $0.001 \sim 0.100$ (本书变异率取为 0.001)。循环往复, 直到进行 100 次迭代为止。从 100 组最优结果中再进行选择, 将适应度最大的数据序列作为最终的优化结果。

2) 模型参数估计

利用反向过程确定数据初值, 结合阻尼最小二乘法, 利用 EViews 统计软件求解出 ARMA 模型满足精度的参数。

对于实际序列 $\{x_t | t=1, 2, \dots, n\}$, 需对其拟合出数学模型:

$$x_t = f(X_t, \beta) + \varepsilon_t \quad (9-49)$$

式中: X_t 为由不同时刻的数据值组成的 k 维向量 $X_t = (x_{t-1}, x_{t-2}, \dots, x_{t-k})^T$; β 为由待估计的模型参数 $\beta_i (i=1, 2, \dots, m)$ 组成的 m 维向量, $\beta = (\beta_1, \beta_2, \dots, \beta_m)^T$; ε_t 是模型的残差; f 表示 X_t 与 β 之间的函数关系; 一般 $k, m < n$ 。

结合 ARMA(p, q) 模型形式, 式 (9-49) 可以记为

$$x_t = X_t \beta + \varepsilon_t$$

式中: $f(X_t, \beta) = X_t \beta$; $X_t = [x_{t-1} \ x_{t-2} \ \cdots \ x_{t-p} \ \varepsilon_{t-1} \ \varepsilon_{t-2} \ \cdots \ \varepsilon_{t-q}]^T$, $p+q=k=m$;
 $\beta = [\phi_1 \ \phi_2 \ \cdots \ \phi_p \ \theta_1 \ \theta_2 \ \cdots \ \theta_q]^T$ 。

由于 $\varepsilon_{t-1}, \varepsilon_{t-2}, \cdots, \varepsilon_{t-q}$ 是 $x_{t-1}, x_{t-2}, \cdots, x_{t-p}$ 的函数, 从而导致 β 的非线性项出现, 故 ARMA 模型 X_t 与 β 之间具有非线性关系, 从而可将 ARMA 模型的目标函数定义为模型的残差平方和:

$$s(\beta) = \sum_{t=p+1}^n \varepsilon_t^2 = \sum_{t=p+1}^n [x_t - f(X_t, \beta)]^2 \quad (9-50)$$

从优化理论的角度来看, 参数 β 的估值问题就是对 $s(\beta)$ 的寻优 (求极小值) 问题。

对于精确求解方程有极大似然法和条件最小二乘法。极大似然法必须是残差序列服从正态分布, 这个条件一般不成立, 所以采用条件最小二乘法。条件最小二乘法中的条件是指初值条件, 当一系列时间序列用 ARMA 模型来拟合, 假设这列序列中第一个时刻之前的 p 个时刻的值为 0, 相对应的残差序列的前 q 个时刻的值为 0。这个条件如果直接用于计算, 会影响参数估计的精确性, 因此采用反向推导过程确定初值, 以提高参数估计的精确性。

假设时间序列 x_1, x_2, \cdots, x_n , 需要去估计 x_{-i} ($i=1, 2, \cdots, n$), 可以由反向模型所产生 $\phi(B)x_k = \theta(B)\phi_k$ 。其中, B 为后移算子, 且有

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) \varepsilon_t \quad (9-51)$$

由于该算子的平稳性, 在一定的范围 $t=l$ ($l \leq n$) 以外, 估计值基本上等于 0。因此, 可以用下式来求解:

$$x_t = \phi^{-1}(B) \theta(B) \phi_t = \sum_{j=0}^{\infty} \psi_j \phi_{t-j} \approx \sum_{j=0}^l \psi_j \phi_{t-j} \quad (9-52)$$

将混合过程用一个 l 阶滑动平均过程代替。接着将模型线性化, 而且要极小化 $S(\phi, \theta) = \sum_{t=1-l}^n [\phi_t]^2$ 。

将 (ϕ_t) 用泰勒级数展开, 使用 β 作为 $m = p + q$ 个参数 (ϕ, θ) 的记号, 在该值附近有一组推测的对应参数值 $\beta_0 = (\beta_{1,0}, \beta_{2,0}, \cdots, \beta_{m,0})$, 则有

$$[\phi_t] = [\phi_{t,0}] - \sum_{i=1}^m (\beta_i - \beta_{i,0}) x_{t,i} \quad (9-53)$$

式中

$$[\phi_{t,0}] = [\phi_t | x, \beta_0], \quad x_{t,i} = -\frac{\partial [a_t]}{\partial \beta_i} \bigg|_{\beta=\beta_0}$$

再结合 ARMA 模型参数的优化理论估计法,采用优化理论的迭代算法寻优,以得到使目标函数 $s(\beta)$ 取值为极小时的模型参数。本书选择使用阻尼最小二乘法进行计算,阻尼最小二乘法是牛顿法与最速下降法的结合,不仅收敛速度快,而且精度高。

对于给定的模型,值 $[\phi_m | x, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{m,0}]$ 对于 $t=1-l, \dots, n$ 用递推方法,然后对 $[\phi_m | x, \beta_{1,0} + \delta_1, \beta_{2,0}, \dots, \beta_{m,0}]$ 重复递推计算,继而又对 $[\phi_m | x, \beta_{1,0}, \beta_{2,0} + \delta_2, \dots, \beta_{m,0}]$ 进行递推计算,由 $x_{t,i} = ([\phi_m | x, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{i,0}, \dots, \beta_{m,0}] - [\phi_m | x, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{i,0} + \delta_i, \dots, \beta_{m,0}]) / \delta_i$ 给出的具有足够精度的所需导数负值,计算出 $[\phi_t]$,最后计算出预期的 ARMA 模型参数。

9.3.3 WGM-ARMA 模型的构建

小波分解与重构是小波变换的关键技术,它是通过不同的带通滤波器将含有综合信息的一组原始时间序列分解成一系列不同特征的时间序列。其中,低频信号反映该序列内在的变化趋势,称为趋势序列;而不同尺度上的高频细节信号则反映随机扰动带来的影响,称为细节序列。分解后的各时间序列所含的频率成分单一,并且小波变换对各序列做了平滑处理,分解后时间序列的平稳性比原时间序列要好得多。灰色系统是建立系统运行趋势模型的有效方法,适用于表示趋势序列的逼近信号预测,所以可以对趋势序列运用灰色模型进行拟合,而对分解后的细节序列采取适合的 ARMA 模型进行拟合。所以,经过小波变换以后再进行预测可以达到比较快的收敛速度和较小的预测误差。

在上述分析和研究的基础上,对导弹装备备件需求预测的思路和框架进行规划。首先利用 Mallat 算法对备件需求序列进行小波分解;然后对趋势序列和细节序列分别进行建模和预测;最后,将各个预测结果进行叠加,便得到原始序列的预测结果。预测框架如图 9-4 所示。

设 $X = \{x(t), t=1, 2, \dots, N\}$ 是备件需求时间序列,用 Mallat 算法对其进行小波分解,可得

$$X = D_1 + D_2 + \dots + D_J + C_J \quad (9-54)$$

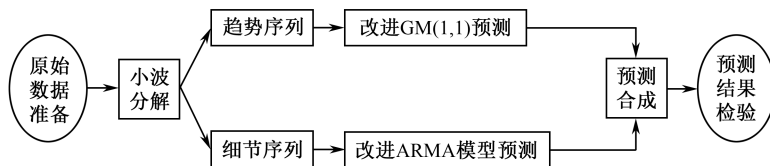


图 9-4 基于 WGM-ARMA 的备件需求预测框架

对重构后的趋势序列 C_j 采用改进的 GM(1,1)模型进行预测；对重构后的细节序列 D_1, D_2, \dots, D_J 采用改进的 ARMA 模型进行拟合。如果已知 $\{x(t), t=1, 2, \dots, k\} (1 \leq k \leq N-1)$ ，确定模型参数后，就可预测 $x(k+1)$ 值，即求

$$x(k+1) = D_{1,k+1} + D_{2,k+1} + \dots + D_{J,k+1} + C_{J,k+1} \quad (9-55)$$

式中： $D_{j,k+1} (j=1, 2, \dots, J)$ 为细节序列预测值； $C_{J,k+1}$ 为趋势序列的预测值。

9.3.4 示例分析

某型导弹是海军舰艇配备的主要反舰武器装备，自从部署到海军舰艇部队以来，多次执行打靶、巡逻、演习等各种战备任务，为维护我国海洋权益发挥了重要作用。通过对海军导弹质量监控数据进行分析可知，该型导弹装备的使用和维修保障活动主要有库存、定检、待发、战备值班、任务返回后性能检测、演习或打靶储运、维修等，由于电流电压不稳或电子器件到寿、振动或冲击、橡胶老化或受力等因素，导致导弹装备出现电子件短路断路或击穿烧蚀、机械件断裂或磨损、橡胶件断裂或磨损等现象，进而导致相应备件的消耗。该型导弹装备某种常用备件 1999—2011 年的需求情况，如表 9-5 所列。

表 9-5 1999—2011 年某种导弹装备备件的需求数据

年份	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
需求量	49	53	60	55	61	65	76	63	72	77	89	80	87

1. 数据分析

由表 9-5 知，备件需求时间序列是非平稳的时间序列，用某一单项预测模型难以取得较满意的效果，因此首先采用小波变换技术把备件需求时间序列进行分解和重构。这里备件需求数据样本共 13 组，选取前 10 组数据样本构建模型，拟合模型；后 3 组数据样本用来作为预测值的比较对象，以检验预测效果的好坏。

2. 需求预测

1) 小波分解

根据 9.1.1 节分析可知, 适合采用 Symlet 小波对备件需求数据组成的时间序列进行了小波分解。根据式 (9-17) ~ 式 (9-26), 可求得需求时间序列的各级小波分解与重构去噪效果各项评价指标值及归一化值, 如表 9-6 所列。从表 9-6 可知, 采用二级小波分解重构时 $H(M)$ 最大, 故最佳分解级数为 2, 即采用 Symlet 小波对备件需求时间序列进行分解。趋势信号、细节信号和重构信号如图 9-4 所示, 重构误差为 0.8153×10^{-10} 。由此可见, 在分解过程中小波基和分解层数的选择是比较合理的。

表 9-6 需求数据的各级分解重构评价指标归一化值

M	1	2	3	4	5
$P_v(M)$	0.8247	1	0.8997	0	0.4679
$P_p(M)$	1	0.8375	0.7616	0.4862	0
$P_{\text{SNR}}(M)$	1	0.6319	0.5140	0.3175	0
$P_r(M)$	0	0.7632	0.9283	0.9893	1
$H(M)$	2.8247	3.2326	3.1036	1.7930	1.4679

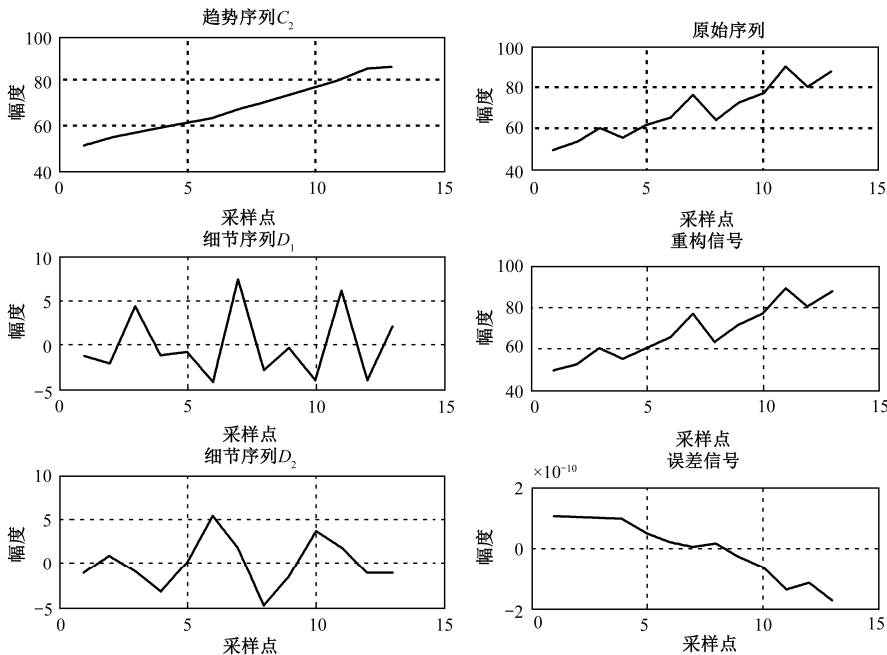
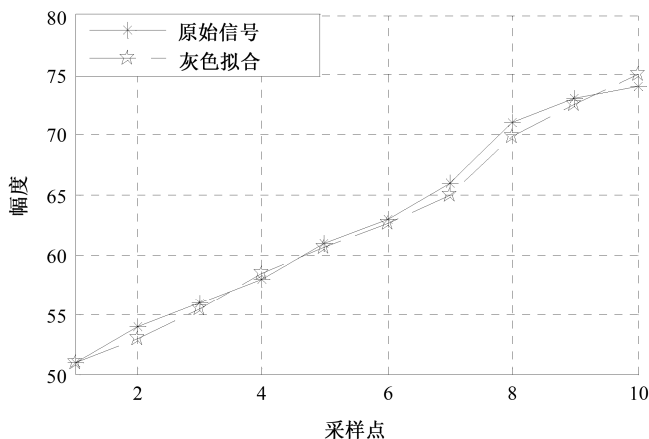


图 9-5 备件需求时间序列的小波分解与重构结果

2) 趋势序列预测

对趋势序列 $\{C_2(i)\} (i=1,2,\dots,13)$ 进行准光滑性检验, 计算得当 $i>2$ 时 $\rho(i)>0.5$, 则 $\{C_2(i)\}$ 是准光滑序列; 再检验 $\{C_2(i)\}$ 的一次累加序列 $\{C_2^{(1)}(i)\}$ 是否具有准指数规律, 计算得 $\sigma^{(1)}(k)<2$, 则 $\{C_2^{(1)}(i)\}$ 具有准指数规律。因此, 可以对趋势序列采用改进的 GM(1,1)模型进行预测, 数据拟合情况如图 9-6 所示。经检验: 平均相对误差 $\bar{\varepsilon}=0.0179$, 关联度 $r=0.9928$, 后验差比值 $C=0.032<0.35$, 小误差概率 $P=1>0.95$, 模型预测精度为一级, 说明可以用于趋势序列预测。

图 9-6 趋势序列 C_2 的灰色拟合

3) 细节序列预测

首先判断细节序列 $\{D_1(i)\} (i=1,2,\dots,13)$ 的平稳性, 如果该序列为非平稳序列, 需要先对其进行差分变换。对 $\{D_1(i)\} (i=1,2,\dots,13)$ 进行 ADF (Augmented Dickey-Fuller) 单位根检验, 结果如表 9-7 所列。在 1% 的显著水平下, 细节序列 $\{D_1(i)\} (i=1,2,\dots,13)$ 拒绝存在一个单位根的原假设, 说明细节序列 $\{D_1(i)\} (i=1,2,\dots,13)$ 是平稳的。选用 AIC 准则进行定阶, 设定 p 和 q 的变化范围为 1~5, 经过多次运算可得到若干个 $ARMA(p,q)$ 模型的 AIC 估计值。比较 AIC 估计值的大小, 得到其最小值为 -4.2762, 并且确定最小值对应的阶数组合为 $(p,q)=(2,1)$ 。IDA-IP-GA 的寻优过程如图 9-7 所示。图 9-7 (a) 是种群平均适应度随遗传代数变化, 图 9-7 (b) 是种群最大适应度随遗传代数变化。可见

随着遗传代数的增加,群体平均适应度在不断增加,最后趋于平稳。寻优曲线中出现了些起伏,这正是 IDA-IP-GA 能避免陷入局部最优解从而产生“早熟”现象的表现。应用阻尼最小二乘法进行参数估计后,将残差序列进行白噪声检验。选择 $k=5$ 作为残差序列自相关系数分析的阶数,分别计算 5 步自相关系数,求得卡方统计量 $F_k=2.519$ 。 F_k 服从自由度为 k 的 χ^2 分布,自由度为 5 的 $\chi^2_{0.05}=6.14$, $F_k < \chi^2_{0.05}$ 。可知残差序列为白噪声序列,ARMA(2,1)可作为小波分解后的细节序列 $\{D_1(i)\}(i=1,2,\cdots,13)$ 的预测模型。细节序列 $\{D_1(i)\}(i=1,2,\cdots,13)$ 拟合情况及其误差相关系数如图 9-8 所示。ARMA 模型的预测信号值和原始信号很接近,预测精度较高,可以用于细节序列的预测。同理,可以对细节序列 $\{D_2(i)\}(i=1,2,\cdots,13)$ 进行分析。

表 9-7 细节序列 D_1 数据 ADF 检验结果

ADF 统计量	显著性水平/%	ADF 检验临界值	P 值
-4.8361	1	-3.6943	0.0007
	5	-2.7517	
	10	-2.4358	

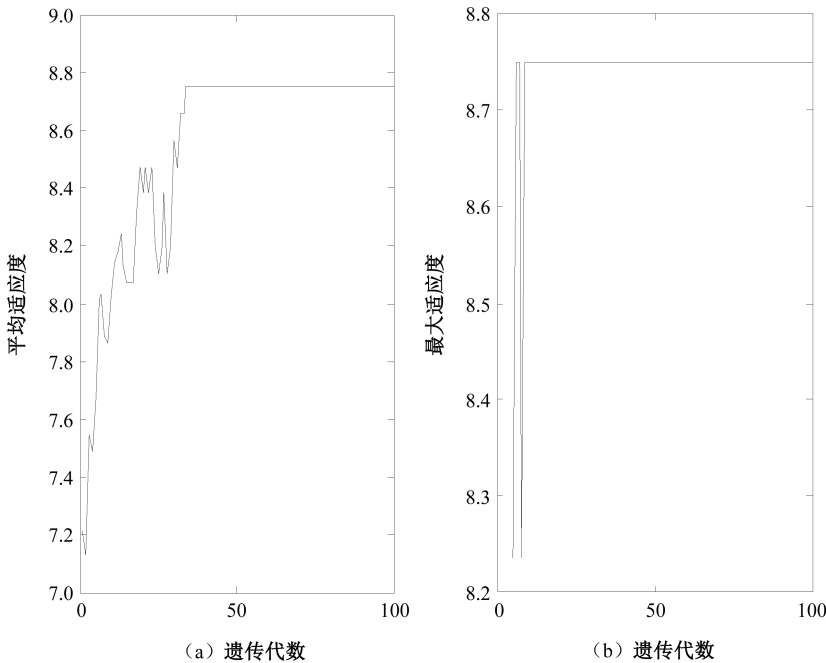
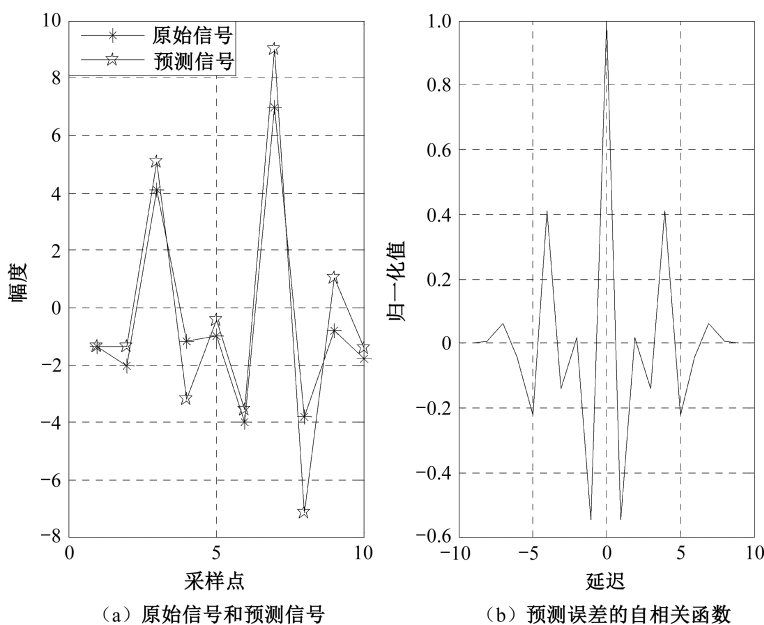


图 9-7 算法寻优过程曲线

图 9-8 细节序列 D_1 的 ARMA 拟合及其误差相关系数

3. 结果分析

在上述分析的基础上,把趋势序列和细节序列的预测值进行叠加,便得到 2009—2011 年的备件需求量预测值。作为对比,再采用传统 GM(1,1)、改进 GM(1,1)、通过 AIC 准则和最小二乘法求解的传统 ARMA 模型、通过改进动态自适应 IP-GA 和阻尼最小二乘法优化的改进 ARMA 模型,分别对备件需求进行预测,这五种方法的预测结果如表 9-8 和图 9-9 所示。选取相对误差(Relative Percent Error, RPE)和平均相对误差(Mean Relative Error, MRE)作为各种预测方法的评价指标。

表 9-8 不同方法的预测值及相对误差

年份	实际值	传统 GM(1,1)		改进 GM(1,1)		传统 ARMA		改进 ARMA		WGM-ARMA	
		预测值	RPE	预测值	RPE	预测值	RPE	预测值	RPE	预测值	RPE
2009	89	94.1329	0.0577	92.4531	0.0388	85.1865	0.0695	92.5972	0.0404	87.5895	0.0158
2010	80	87.2335	0.0904	85.7892	0.0724	83.7340	0.0467	82.7541	0.0344	82.1363	0.0267
2011	87	90.3587	0.0386	89.4527	0.0282	89.8592	0.0328	89.4634	0.0283	86.7269	0.0031
MRE		0.0622		0.0465		0.0497		0.0344		0.0152	

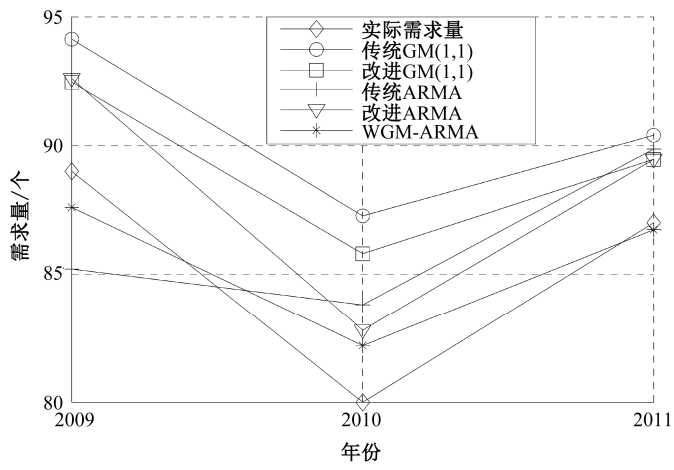


图 9-9 预测结果对比

由表 9-8 可知，改进 GM(1,1)模型的预测性能高于传统 GM(1,1)模型，基于改进动态自适应 IP-GA 和阻尼最小二乘法优化的 ARMA 模型的预测精度高于选用 AIC 准则和最小二乘法优化的 ARMA 模型，基于小波变换和改进 GM(1,1)-ARMA 的预测方法预测准确度高于其他预测方法，而且性能稳定。

9.4 基于改进 Theil 不等系数的备件需求预测方法

传统的组合预测模型是从改善某种拟合误差角度建立的，同一种单项预测方法在各个时点的权系数是一样的。然而这与实际情况并不相符，因为同一种单项预测方法在不同的时刻表现并不是完全相同。因此，本书建立基于 Theil 不等系数的 IOWHA 算子的组合预测模型，它不直接考虑预测误差的大小，而是从相关性指标的角度进行组合预测模型的构建。另外，在各个时间点上每个模型的权重系数都与该时点的拟合精度有关，因此，能较好地反映备件需求实际情况。但在实际预测时预测期数据是未知的，无法直接利用该方法进行预测。目前，通常的做法是用近几期平均精度高的单项预测方法来反映预测区间未来几期的预测精度，但缺乏理论基础且预测效果不佳。本书利用马尔可夫链表述待预测时点的预测精度状态的方法，并提出通过改进免疫记忆万有引力搜索算法（IMGSA）求解最优权系数。

9.4.1 单项预测模型的建立

1. 基于最小二乘法优化的指数平滑法

指数平滑法作为时间序列分析中一种常见的方法,适用于数据较少、有周期性和趋势性变动情况下的预测。

1) 平滑指数法概述

设 α 为平滑系数, $\{X_i | i = 0, 1, 2, \dots, t\}$ 为时间序列, $\{\bar{X}_i | i = 1, 2, \dots, t\}$ 为平滑后的拟合序列, \bar{X}_0 为平滑初始值。则指数平滑模型为

$$\bar{X}_t = \alpha X_t + (1 - \alpha) \bar{X}_{t-1} = \sum_{i=0}^{t-1} \alpha (1 - \alpha)^i X_{t-i} + (1 - \alpha)^t \bar{X}_0, \quad (t = 1, 2, \dots, n) \quad (9-56)$$

2) 平滑系数的优化

为了克服绝对误差最小二乘法评价一个近似值精确度的不足,本书采用相对误差最小二乘法来保证估计的精度,解决了人为随机选择时误差大和难以修正的缺陷。其优化确定 α 的过程如下:

令目标函数为 G , 则有

$$G = \sum_{i=1}^n \left[\frac{\bar{X}_i - X_i}{\bar{X}_i} \right]^2 = \sum_{i=1}^n \left[\left(\sum_{t=0}^{i-1} \alpha (1 - \alpha)^t X_{i-t} + (1 - \alpha)^i X_0 - X_i \right) / \bar{X}_i \right]^2 \quad (9-57)$$

$$\frac{dG}{d\alpha} = \sum_{i=1}^n \left\{ 2 \left[\left(\sum_{t=0}^{i-1} \alpha (1 - \alpha)^t X_{i-t} + (1 - \alpha)^i X_0 - X_i \right) / \bar{X}_i \right] \cdot \left[\sum_{t=0}^{i-1} (1 - \alpha)^{t-1} (1 - \alpha - \alpha t) \frac{X_{i-t}}{X_i} - \frac{i(1 - \alpha)^{i-1} X_0}{X_i} \right] \right\} \quad (9-58)$$

为了使 G 取得最小值, 令 $\frac{dG}{d\alpha} = 0$, 则结果如下:

$$\sum_{i=1}^n \left\{ \left[\sum_{t=0}^{i-1} \alpha X_{i-t} / \bar{X}_i + \left(\frac{X_0}{\bar{X}_i} - 1 \right) \right] \cdot \left[\sum_{t=0}^{i-1} \frac{(1 - t) X_{i-t}}{X_i} - \sum_{t=0}^{i-1} \frac{t X_{i-t}}{(1 - \alpha) X_i} \right] \right\} = 0 \quad (9-59)$$

为计算简便, 令

$$\sum_{t=0}^{i-1} X_{i-t} / \bar{X}_i = P, \quad \sum_{t=0}^{i-1} \frac{(1 - t) X_{i-t}}{X_i} = Q, \quad \sum_{t=0}^{i-1} \frac{t X_{i-t}}{X_i} = R, \quad \frac{X_0}{\bar{X}_i} - 1 = S$$

则式 (9-59) 简化为

$$\sum_{i=1}^n \left[(\alpha P + S) \left(Q - \frac{R}{1 - \alpha} \right) \right] = 0 \quad (9-60)$$

解式 (9-60) 得到 α 的优化值为

$$\alpha^* = \frac{\sum_{i=1}^n PQ - \sum_{i=1}^n PR - \sum_{i=1}^n SQ \pm \sqrt{\left(\sum_{i=1}^n PQ - \sum_{i=1}^n PR - \sum_{i=1}^n SQ\right)^2 - 4 \sum_{i=1}^n PQ \cdot \left(\sum_{i=1}^n SR - \sum_{i=1}^n SQ\right)}}{2 \sum_{i=1}^n PQ} \quad (9-61)$$

3) 初始值的确定

为克服平滑指数初始值确定依靠经验的不足, 用最小二乘法对前几个数进行拟合, 估计 a_0 和 b_0 , 再根据 a_0 和 b_0 的关系式计算初始值。以二次指数平滑模型 $\bar{X}_{t+T} = a_t + b_t \cdot T$ ($T=1, 2, \dots$) 参数的估计过程为例, 确定初始值的过程如下:

$$a_t = 2\bar{X}_t^{(1)} - \bar{X}_t^{(2)}, \quad b_t = \frac{\alpha}{1-\alpha}(\bar{X}_t^{(1)} - \bar{X}_t^{(2)}) \quad (9-62)$$

式中: $\bar{X}_t^{(1)}$ 、 $\bar{X}_t^{(2)}$ 为当前时期 T 时的一次指数平滑值和二次指数平滑值。

解得 $\bar{X}_t^{(1)}$ 、 $\bar{X}_t^{(2)}$, 并将 $t=0$ 代入得

$$\bar{X}_0^{(1)} = a_0 - \frac{1-\alpha}{\alpha} b_0, \quad \bar{X}_0^{(2)} = a_0 - \frac{2(1-\alpha)}{\alpha} b_0 \quad (9-63)$$

假设 X_1 、 X_2 满足线性平滑方程, 则有 $X_1 = a_0$, $X_2 = a_0 + b_0$, 即

$$a_0 = X_1, \quad b_0 = X_2 - X_1 \quad (9-64)$$

将式 (9-64) 代入 $\bar{X}_0^{(1)}$ 和 $\bar{X}_0^{(2)}$, 即可计算出其初始值。

2. 基于数据变换技术改进的 GM(1,1)

GM(1,1) 对时间序列短、信息不完全、统计数据少的建模具有独特功能。当数据样本是波动性较强的振荡序列时, GM(1,1) 模型在数据拟合时不能有效地体现数据样本的变化趋势, 且有时不能满足 GM(1,1) 的建模条件。因此, 采用数据变换技术对 GM(1,1) 模型进行改进, 力求从根本上改变原始数据的波动性。

1) 振荡序列的改进

定义 9.1 设原始数据序列 $X = (x(1), x(2), \dots, x(n))$, 若 $\exists k, k' \in \{2, 3, \dots, n\}$, 有 $x(k) - x(k-1) > 0$, $x(k') - x(k'-1) < 0$, 则称 X 为振荡序列。

定义 9.2 设原始数据序列 $X = (x(1), x(2), \dots, x(n))$, 记 $M = \max\{x(k) | k=1, 2, \dots, n\}$, $m = \min\{x(k) | k=1, 2, \dots, n\}$, $T = M - m$, 则称变换 $x(k)d_1 = x(k) + (k-1)T$ ($k=1, 2, \dots, n$) 为平移变换, 记作 D_1 。

定义 9.3 设原始数据序列 $X = (x(1), x(2), \dots, x(n))$ ($x(i) > 0; i=1, 2, \dots, n$)

则称变换 $x(k)d_2 = \left(\prod_{i=1}^k x(i) \right)^{\frac{1}{k}}$ ($k=1,2,\dots,n$) 为几何平均变换, 记为 D_2 。

经论证可知: 平移变换可以使任意的原始数据序列 X 变为一个单调递增的序列 X_1 ; 几何平均变换不但可使序列 X_1 的随机波动性减弱, 消除冲击干扰项对系统造成的数据“失真”现象, 而且能保持原有数据序列 X_1 的单调性, 并提高序列 X_1 的光滑性。因此, 通过对平移变换和几何平均变换后生成得到的数据序列建立 GM(1,1)模型, 可以有效地提高预测模型的拟合精度。

2) 模型的建立

设 $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$ 为振荡序列, 先对 $X^{(0)}$ 进行平移变换, 所得序列记作 $X_1^{(0)} = \{x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(n)\}$, $x_1^{(0)}(k) = x^{(0)}(k)d_1$, $k=1,2,\dots,n$ 。然后对 $X_1^{(0)}$ 进行几何平均变换, 新序列记作 $Y^{(0)} = \{y^{(0)}(1), y^{(0)}(2), \dots, y^{(0)}(n)\}$, 其中, $y^{(0)}(k) = x_1^{(0)}(k)d_2$ ($k=1,2,\dots,n$)。以 $Y^{(0)} = \{y^{(0)}(1), y^{(0)}(2), \dots, y^{(0)}(n)\}$ 为原始序列建立 GM(1,1)微分方程模型:

$$\begin{cases} \frac{dy^{(1)}(t)}{dt} + ay^{(1)}(t) = b \\ y^{(1)}(1) = y^{(0)}(1) \end{cases} \quad (9-65)$$

求解上述微分方程的响应函数:

$$\hat{y}^{(1)}(k) = \left(y^{(1)}(1) + c - \frac{b}{a} \right) e^{-a(k-1)} + \frac{b}{a} \quad (k=1,2,\dots,n) \quad (9-66)$$

做一次累减生成, 可得

$$\hat{y}^{(0)}(k+1) = (1 - e^a) \left[y^{(0)}(1) + c - \frac{b}{a} \right] e^{-ak} \quad (9-67)$$

对 $\{\hat{y}^{(0)}(k)\} = \{\hat{y}^{(0)}(1), \hat{y}^{(0)}(2), \dots, \hat{y}^{(0)}(n)\}$ 进行还原, 可得

$$\{\hat{x}_1^{(0)}(k)\} = \{\hat{x}_1^{(0)}(1), \hat{x}_1^{(0)}(2), \dots, \hat{x}_1^{(0)}(n)\}$$

由 $\hat{y}^{(0)}(k) = \left[\prod_{i=1}^k \hat{x}^{(0)}(i) \right]^{\frac{1}{k}}$ 可得

$$\{\hat{x}_1^{(0)}(k)\} = \frac{[\hat{y}^{(0)}(k)]^k}{\prod_{i=1}^{k-1} \hat{x}^{(0)}(i)} \quad (k=2,3,\dots,n) \quad (9-68)$$

式中: $\hat{x}_1^{(0)}(1) = \hat{y}^{(0)}(1)$ 。

从而可得

$$\{\hat{x}_1^{(0)}(k)\} = \frac{\hat{y}^{(0)}(k)^k}{\hat{y}^{(0)}(k-1)^{k-1}} (k=2,3,\dots,n) \quad (9-69)$$

式中: $\hat{x}_1^{(0)}(1) = \hat{y}^{(0)}(1)$ 。

进一步还原可得

$$\{\hat{x}^{(0)}(k)\} = \hat{x}_1^{(0)}(k) - (k-1)T (k=1,2,\dots,n) \quad (9-70)$$

式中: $\hat{x}^{(0)}(1) = \hat{x}_1^{(0)}(1)$ 。

3. 改进的 WLS-SVM

加权最小二乘支持向量机 (Weighted Least Squares Support Vector Machine, WLS-SVM) 的回归算法, 是通过非线性变换在高维空间中寻找输入向量和输出变量之间的非线性关系, 能够很好地反映备件需求影响因素的作用。

WLS-SVM 是对 LS-SVM 进行了改进, 根据训练误差给样本分配权值, 让含有较大误差的点对模型的贡献度低, 使得模型具有鲁棒性。

对于训练数据集 $T = \{x_i \in \mathbf{R}^n, y_i \in \mathbf{R} | i=1,2,\dots,n\}$, WLS-SVM 的函数估计问题可以描述为求解

$$\begin{aligned} \min_{w,b,e} J(w,e) &= \frac{1}{2} \|w\|^2 + \gamma \frac{1}{2n} \sum_{i=1}^n v_i e_i^2 \\ \text{s.t. } y_i &= w^T \phi(x_i) + b + e_i (i=1,2,\dots,n) \end{aligned} \quad (9-71)$$

式中: J 为损失函数 (结构风险); w 为原始权空间的权值向量; b 为偏置向量, $e = (e_1, e_2, \dots, e_n)^T$ 为残差向量, $e_i \in \mathbf{R}$ 为残差; v_i 为残差 e_i 的加权因子; $\gamma \in \mathbf{R}_+$ 为惩罚系数 (容量调节参数); $\phi(\cdot)$ 为将输入空间数据映射到高维特征空间的非线性映射函数。

本书中提出的加权方法源于偏鲁棒 M-回归 (Partial Robust M-regression, PRM), 并做了一定的改进, 取消了原方法中迭代修正得分向量的过程, 使之能够应用到 WLS-SVM。加权方法具体地将离群点分成两类: 一类是高杠杆点, 它们是远离输入数据中心的样本点; 另一类是高残差点, 是输出预测值与实际值相差较大的样本点。通过此加权方法不仅可以降低高残差点对模型的干扰, 也可以将高杠杆点的不利影响降到最低。

权值确定方法如下:

设第 i 个样本数据杠杆权值 w_i^x ，即

$$w_i^x = f \left[\frac{\|x_i - \text{med}(X)\|}{\text{med}(\|x_i - \text{med}(X)\|)}, c \right] \quad (i=1, 2, \dots, n) \quad (9-72)$$

且

$$f(z, c) = \frac{1}{(1 + |z/c|)^2} \quad (9-73)$$

式中： $\|\cdot\|$ 为欧几里得距离； $\text{med}(\cdot)$ 为中位值； x_i 为第 i 个样本数据； c 为常数（通常取 4）； X 为所有样本。

设第 i 个样本数据的残差权值 w_i^e ，即

$$w_i^e = f \left[\frac{e_i}{\tilde{e}}, c \right] \quad (9-74)$$

式中： e_i 为第 i 个样本数据预测值与实际值的残差； \tilde{e} 为残差的鲁棒尺度估计，且有

$$\tilde{e} = \text{med}_i |e_i - \text{med}_j(e_j)| \quad (i, j=1, 2, \dots, n) \quad (9-75)$$

综合考虑上述两种权值，则第 i 个样本的权值为

$$w_i = \sqrt{w_i^x w_i^e} \quad (9-76)$$

9.4.2 基于 Theil 不等系数的 IOWHA 算子的组合预测模型构建

1. IOWHA 算子定义

定义 9.4 设 $\langle u_1, a_1 \rangle, \langle u_2, a_2 \rangle, \dots, \langle u_m, a_m \rangle$ 为二维数组，令

$$g_w(\langle u_1, a_1 \rangle, \langle u_2, a_2 \rangle, \dots, \langle u_m, a_m \rangle) = 1 / \sum_{i=1}^m \frac{w_i}{a_{u-\text{index}(i)}} \quad (9-77)$$

则称函数 g_w 为 u_1, u_2, \dots, u_m 所产生的 m 维诱导有序加权调和平均算子(IOWHA)。式中， u_i 为 a_i 的诱导值， $u-\text{index}(i)$ 为 u_1, u_2, \dots, u_m 中按从大到小的顺序排列的第 i 个大的数的下标； $w = (w_1, w_2, \dots, w_m)^T$ 是与 g_w 有关的加权向量，满足

$$\sum_{i=1}^m w_i = 1 (w_i \geq 0; i=1, 2, \dots, m)。$$

定义 9.4 表明，IOWHA 算子是对诱导值 u_1, u_2, \dots, u_m 按从大到小的顺序排序后所对应的 a_1, a_2, \dots, a_m 中的数进行有序加权几何平均，权系数 w_i 与数 a_i 的大小和位置无关，而是与其诱导值所在的位置有关。

2. Theil 不等系数定义

定义 9.5

令

$$\tau = \sqrt{\sum_{t=1}^N (x_t - \hat{x}_t)^2} / \sqrt{\sum_{t=1}^N (x_t)^2} \quad (9-78)$$

则称 τ 为组合预测值 $\{\hat{x}_t | t=1, 2, \dots, N\}$ 序列与实际观察值 $\{x_t | t=1, 2, \dots, N\}$ 序列的 Theil 不等系数。 τ 值越小, 表示组合预测精度越高; 当 $\tau=0$ 时, $\hat{x}_t = x_t (t=1, 2, \dots, N)$, 表示组合预测准确无误。

3. 组合预测模型建立

设备件年需求量的时间序列为 $\{x_t, t=1, 2, \dots, N\}$, 组合模型中共有 m 种单项预测方法, 其中, 第 i 种预测方法第 t 时刻的预测值记为 $x_{it} (i=1, 2, \dots, m; t=1, 2, \dots, N)$ 。同时设 l_1, l_2, \dots, l_m 为组合模型的权系数。它满足归一性和非负性, 即

$$\sum_{i=1}^m l_i = 1 (l_i \geq 0; i=1, 2, \dots, m)$$

建模步骤如下:

(1) 计算第 i 种单项预测方法第 t 时刻的预测精度, 即

$$a_{it} = \begin{cases} 1 - |(x_t - x_{it})/x_t|, & |(x_t - x_{it})/x_t| \leq 1 \\ 0, & |(x_t - x_{it})/x_t| > 1 \end{cases} \quad (i=1, 2, \dots, m; t=1, 2, \dots, N) \quad (9-79)$$

显然 $a_{it} \in [0, 1]$, 将预测精度 a_{it} 看成预测值 x_{it} 的诱导值, 从而得到组合模型中第 t 时刻各单项预测方法的预测精度值和其对应的在样本区间上的预测值的二维数组 $\langle a_{1t}, x_{1t} \rangle, \langle a_{2t}, x_{2t} \rangle, \dots, \langle a_{mt}, x_{mt} \rangle$ 。

(2) 设 $a\text{-index}(it)$ 是 $a_{1t}, a_{2t}, \dots, a_{mt}$ 第 i 个大的预测精度的下标, 计算 m 种单项预测方法在第 t 时刻的 IOWHA 的组合预测值, 即

$$\hat{x}_t = g_t(\langle a_{1t}, x_{1t} \rangle, \langle a_{2t}, x_{2t} \rangle, \dots, \langle a_{mt}, x_{mt} \rangle) = \frac{1}{\sum_{i=1}^m \frac{l_i}{x_{a\text{-index}(it)}}} \quad (9-80)$$

式 (9-80) 表明, 组合预测的赋权系数与单项预测方法的预测值大小无关, 而是与单项预测方法在各时点上的预测精度的大小密切相关。

(3) 计算第 t 时刻组合预测模型的误差。令

$$e_{it} = \frac{1}{x_t} - \frac{1}{x_{it}}, \quad e_t = \frac{1}{x_t} - \frac{1}{\hat{x}_t}$$

式中: e_{it} 为第 i 种单项预测方法第 t 时刻预测值与相应的实际值之间的倒数误差; e_t 为第 t 时刻的组合预测值与相应的实际值之间的倒数误差。则有

$$e_i = \frac{1}{x_t} - \frac{1}{\hat{x}_t} = \frac{1}{x_t} - \sum_{i=1}^m l_i \frac{1}{x_{a-\text{index}(it)}} = \sum_{i=1}^m l_i \left[\frac{1}{x_t} - \frac{1}{x_{a-\text{index}(it)}} \right] = \sum_{i=1}^m l_i e_{a-\text{index}(it)} \quad (9-81)$$

式中

$$e_{a-\text{index}(it)} = \frac{1}{x_t} - \frac{1}{x_{a-\text{index}(it)}}$$

(4) 计算 m 阶 IOWHA 的组合预测协方差信息方阵:

$$\begin{aligned} \sum_{t=1}^N \left(\frac{1}{x_t} - \frac{1}{\hat{x}_t} \right)^2 &= \sum_{t=1}^N \left(\sum_{i=1}^m l_i e_{a-\text{index}(it)} \right)^2 = \sum_{t=1}^N \left(\sum_{i=1}^m \sum_{j=1}^m l_i l_j e_{a-\text{index}(it)} e_{a-\text{index}(jt)} \right) \\ &= \sum_{i=1}^m \sum_{j=1}^m l_i l_j \left(\sum_{t=1}^N e_{a-\text{index}(it)} e_{a-\text{index}(jt)} \right) = \sum_{i=1}^m \sum_{j=1}^m l_i l_j E_{ij} = \mathbf{L}^T \mathbf{E} \mathbf{L} \end{aligned} \quad (9-82)$$

式中: $E_{ij} = \sum_{t=1}^N e_{a-\text{index}(it)} e_{a-\text{index}(jt)}$ ($i, j = 1, 2, \dots, m$); $\mathbf{L} = (l_1, l_2, \dots, l_m)^T$; $\mathbf{E} = (E_{ij})_{m \times m}$

为 m 阶 IOWHA 算子的组合预测协方差信息方阵。

(5) 计算 \hat{x}_t 与 x_t 的 Theil 不等系数。令

$$\begin{cases} \tau_i = \sqrt{\sum_{t=1}^N \left(\frac{1}{x_t} - \frac{1}{x_{it}} \right)^2} / \sqrt{\sum_{t=1}^N \left(\frac{1}{x_t} \right)^2} \\ \tau = \sqrt{\sum_{t=1}^N \left(\frac{1}{x_t} - \frac{1}{\hat{x}_t} \right)^2} / \sqrt{\sum_{t=1}^N \left(\frac{1}{x_t} \right)^2} \end{cases} \quad (1, 2, \dots, m) \quad (9-83)$$

则称 τ_i 和 τ 分别为第 i 种单项预测方法预测值倒数序列和基于 IOWHA 算子组合预测值倒数序列与实际值倒数序列的 Theil 不等系数。显然 Theil 不等系数 $\tau_i \in (0, +\infty)$, 且 $\tau \in (0, +\infty)$, τ 值越小, 表示组合预测精度越高。则有

$$\tau = \sqrt{\mathbf{L}^T \mathbf{E} \mathbf{L}} / \sqrt{\sum_{t=1}^N \left(\frac{1}{x_t} \right)^2} \quad (9-84)$$

式 (9-84) 表明, 基于 Theil 不等系数的 IOWHA 算子组合预测值序列与实际值序列的 Theil 不等系数为组合预测方法的权系数 l_1, l_2, \dots, l_m 的函数, 记为 $\tau(l_1, l_2, \dots, l_m)$ 。 $\tau(l_1, l_2, \dots, l_m)$ 越小, 表示组合预测方法越有效。

(6) 建立基于 Theil 不等系数的 IOWHA 算子组合预测模型:

$$\begin{aligned} \min \tau(l_1, l_2, \dots, l_m) &= \sqrt{\mathbf{L}^T \mathbf{E} \mathbf{L}} / \sqrt{\sum_{t=1}^N \left(\frac{1}{x_t} \right)^2} \\ \text{s.t. } \begin{cases} \sum_{i=1}^m l_i = 1 \\ l_i \geq 0, i = 1, 2, \dots, m \end{cases} \end{aligned} \quad (9-85)$$

定义 9.6 记 $\tau_{\min} = \min_{1 \leq i \leq m} \tau_i$, $\tau_{\max} = \max_{1 \leq i \leq m} \tau_i$: 若 $\tau(l_1, l_2, \dots, l_m) < \tau_{\min}$, 则称组合预测模型 (9-85) 为优性组合预测; 若 $\tau_{\min} \leq \tau(l_1, l_2, \dots, l_m) \leq \tau_{\max}$, 则称组合预测模型 (9-85) 为非劣性组合预测; 若 $\tau(l_1, l_2, \dots, l_m) > \tau_{\max}$, 则称组合预测模型 (9-85) 为劣性组合预测。

定义 9.6 表明只有 IOWHA 组合预测方法的预测值倒数序列与指标值倒数序列的 Theil 不等系数小于各单项预测值倒数序列与指标实际值的倒数序列的 Theil 不等系数中最小者, 该组合预测模型才是优性的。

9.4.3 基于 Theil 不等系数的 IOWHA 算子的组合预测模型改进

1. 基于改进万有引力搜索算法的模型优化求解

1) 万有引力搜索算法的基本原理

万有引力搜索算法 (Gravitational Search Algorithm, GSA) 是由 E.Rashedi 等在 2009 年提出的一种源于对物理学中的万有引力进行模拟得到的优化搜索技术, 是一种元启发式算法。它通过群体中各物体之间的万有引力相互作用产生的群体智能指导优化搜索。

2) 生物免疫系统及其特性

生物免疫系统是由器官、细胞和分子组成的一个复杂系统, 它是除神经系统外, 机体能识别“非己”和“自己”的刺激, 对之做出精确的应答, 并保留记忆反应的功能系统。研究表明, 生物免疫系统具有多种特性, 如产生抗体多样性的能力、自我调节机制、免疫记忆功能等, 免疫系统的这些特性已经引起工程研究人员的关注, 就免疫学原理人们已提出了不少新的基于免疫特性的优化方法、算法和技术。从生物信息处理的角度来看, 免疫系统具有多样性、免疫自我调节和免疫记忆特性的特性。

3) 基于免疫记忆和调节机制的改进万有引力搜索算法

试验证明, GSA 算法存在早熟收敛、易陷入局部最优、缺少有效加速机制的不足。为了解决以上不足, 把免疫系统的免疫信息处理机制 (多样性、免疫自我调节、免疫记忆等) 引入万有引力搜索算法中, 在万有引力搜索算法的基本框架上, 结合免疫系统的免疫信息处理机制, 提出免疫记忆万有引力搜索算法。这种免疫记忆万有引力搜索算法结合了万有引力搜索算法具有的全局寻优能力和免疫系统的免疫信息处理机制, 从而改进了万有引力搜索算法易于陷入局部极值点的不足, 提高了进化后期算法的收敛速度和精度。

抗体多样性和免疫记忆是免疫系统的重要特性,在免疫调节中,与抗原亲和力大并且浓度较低的抗体会受到促进,而与抗原亲和力小或浓度较高的抗体将会受到抑制,以此保证抗体的多样性;免疫记忆是免疫系统将与入侵抗原反应部分的抗体作为记忆细胞保留下来,对于同类抗原的再次侵入,相应的记忆细胞被激活而产生大量的抗体。把这种多样性和免疫记忆特性引入万有引力搜索算法中,可提高算法的全局搜索能力而不致陷于局部解。抗体的浓度越小,选择的概率越大;抗体的浓度越大,选择概率越小。这样,在保留高适应度个体的同时,进一步确保了抗体的多样性,能避免早熟现象。

在抗体群更新过程中,总是希望适应度高的抗体被保留下来,但如果此类抗体过于集中,即浓度过高,则很难保证抗体的多样性,很容易使算法陷入局部极优,而丢失那些适应度较差却保持着较好进化趋势的抗体。因此,采用了基于浓度机制的多样性保持策略,使得新一代抗体群中,各适应度层次的抗体维持着一定的浓度。第 i 个抗体的浓度为

$$D(x_i) = \frac{1}{\sum_{j=1}^{N+M} |f(x_i) - f(x_j)|} \quad (i=1,2,\dots,N+M) \quad (9-86)$$

由式 (9-86) 可推导出基于抗体浓度的概率为

$$P(x_i) = \frac{\frac{1}{D(x_i)}}{\sum_{i=1}^{N+M} \frac{1}{D(x_i)}} = \frac{\sum_{j=1}^{N+M} |f(x_i) - f(x_j)|}{\sum_{i=1}^{N+M} \sum_{j=1}^{N+M} |f(x_i) - f(x_j)|} \quad (i=1,2,\dots,N+M) \quad (9-87)$$

式中: x_i 和 $f(x_i)$ ($i=1,2,\dots,N+M$) 分别为第 i 个抗体及其适应度函数值。

由式 (9-87) 可知,与抗体 i 相似的抗体越多,抗体 i 被选中的概率就越小;与抗体 i 相似的抗体越少,抗体 i 被选中的概率就越大。这使得低适应度个体也可获得进化的机会。因此,基于抗体浓度的概率选择公式在理论上保证了抗体的多样性。

算法流程如下:

(1) 初始化:针对待求问题的特征,从记忆库中搜寻该类问题的记忆抗体或随机产生 N 个抗体 x_i 及其“运动”速度 v_i 形成初始抗体群 P_0 。

(2) 生成免疫记忆抗体:计算当前抗体群 P_k 中抗体的适应度值及 p_j ($j=1,2,\dots,N$),并把 p_g 作为免疫记忆抗体存入记忆库中,判断是否满足结束条件。若满足结束条件,则停止运行并输出结果;否则,继续。

(3) 新抗体的生成:一是根据抗体的速度更新原则产生 N 个新抗体;二是

随机产生 M 个新抗体。

(4) 基于浓度的抗体选择：用式 (9-87) 计算步骤 (3) 中生成的 $N+M$ 个新抗体的选择概率，依概率大小选择 N 个抗体形成抗体群 A_k 。

(5) 抗体群更新：将记忆库中的免疫记忆抗体替换抗体群 A_k 中适应度较差的若干抗体，形成新一代抗体群 P_{k+1} ，然后，转步骤 (2)。

这一改进算法称为免疫记忆万有引力搜索算法 (Immune Memory Gravitational Search Algorithm, IMGSA)。其求解流程如图 9-10 所示。

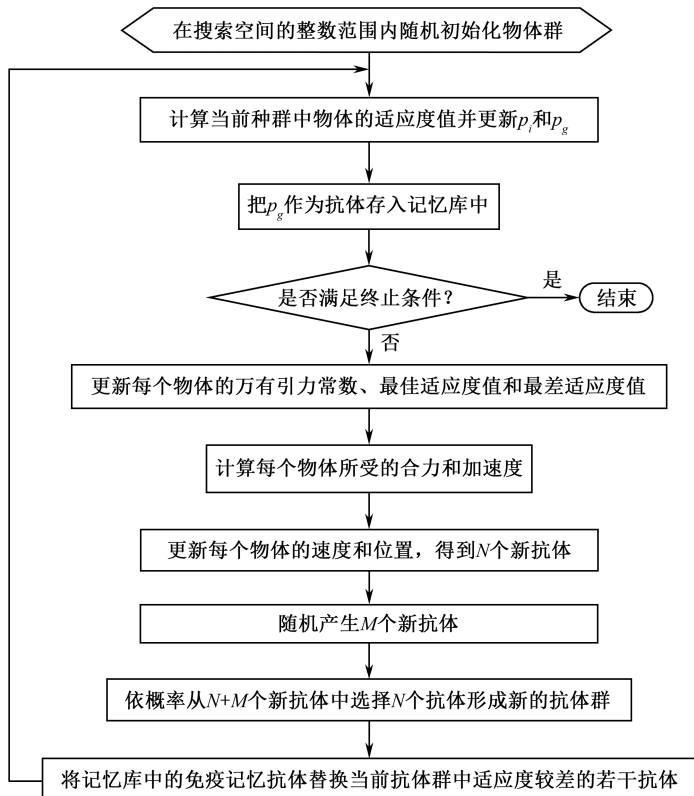


图 9-10 算法求解流程

4) 适应度函数的构造

式 (9-85) 建立的组合预测模型是一个典型的有约束优化问题，在求解有约束优化问题时，常采用罚函数法。罚函数法的基本思想是：把一般较复杂的有约束最优化问题转化为简单的无约束问题，通过求出无约束问题的最优解来逼近原有约束最优化问题的解。

具体做法是：根据约束的特点，设法对个体违背约束条件的情况给予惩罚，构造某种惩罚函数；然后把它加到目标函数中，将约束优化问题的求解转化为附带考虑代价或惩罚的无约束问题的求解。这种惩罚策略对于在无约束问题求解过程中企图违反约束的迭代点给予很大的目标函数值，迫使一系列无约束问题的极小点或者无限地靠近可行域，或者一直保持在可行域内移动，直至迭代点收敛到原约束问题的极小点。根据组合预测模型的建立过程，构造扩展目标函数如下：

$$\text{fitness} = \min \left(\sum_{i=1}^m l_i e_{a-\text{index}(it)} \right)^2 + M [\min(0, l_i)] + N \{ \min[0, (1-l_i)] \} \quad (9-88)$$

式中： M 和 N 为惩罚因子，是两个充分大的正数。实际计算中，惩罚因子 M 和 N 的取值太小起不到惩罚作用，太大则由于误差的影响会导致错误，因此要尽量保证 M 和 N 与粗略估计的最小值在一个数量级上。

2. 马尔可夫链理论概述

马尔可夫过程是无后效性的随机过程。当状态和时间参数都为离散时，称为马尔可夫链。根据导弹装备备件需求时间序列的特点，可以将其近似看作状态和时间参数都是离散的马尔可夫过程。

设某系统在时刻 $t=n$ 有 k 个可能状态，即 $X_n=1, 2, \dots, k (n=0, 1, \dots)$ ， $a_i(n)$ 表示系统在时刻 $t=n$ 处于状态 $i (i=1, 2, \dots, k)$ 的概率，即 $a_i(n)=P(X_n=i)$ ，且 $\sum_{i=1}^k a_i(n)=1$ 。系统在时刻 n 状态为 i ，时刻 $n+1$ 转移到状态 j 的概率为 $P_{ij}(i, j=1, 2, \dots, k)$ ，即 $P_{ij}=P(X_{n+1}=j|X_n=i)$ ，称 P_{ij} 为一步转移概率， $\mathbf{P}=\{P_{ij}\}$ 为一步转移概率矩阵，其中 $P_{ij} \geq 0$ ，且 $\sum_{j=1}^k P_{ij}=1$ 。

若系统在 $t=n$ 时间处于状态 i ，经过 k 次转移后在 $t=n+k$ 时间处于状态 j 的概率为 $P_{ij}(k)$ ，则称 $P_{ij}(k)$ 为 k 步转移概率，记作 $\mathbf{P}^{(k)}$ 。不难证明，它可由一步转移概率矩阵来表示：

$$\mathbf{P}^{(k)} = \mathbf{P}^k \quad (9-89)$$

如果起始状态概率为 $\mathbf{P}^{(0)}$ ，经过 k 步转移后处于 j 状态的概率为

$$\mathbf{P}^{(k)} = \mathbf{P}^{(k-1)} \cdot \mathbf{P}^{(1)} = \mathbf{P}^{(0)} \mathbf{P}^{(k)} \quad (9-90)$$

3. 建模过程及步骤

在上述分析的基础上，应用马尔可夫链原理和IMGSA对建立的组合预测模

型进行改进，具体如下：

(1) 对于已知样本点区间，应用 IMGSA 求解式 (9-85) 组合模型中 m 种单项预测模型的权系数 l_1, l_2, \dots, l_m 。

(2) 对于未来预测点区间，以预测精度为基准，划分若干状态空间。由组合模型中各单项预测模型预测精度的最大值和最值得到其变化区间，然后按照聚类分析思想，认为预测精度相近的点为同一状态，进而划定合适的步长后，就可以得到各单项模型在各时点预测精度的状态。

(3) 根据预测期数 T 的不同，按照预测要求，构造不同的 T 步状态转移概率矩阵，即 P^1, P^2, \dots, P^T 。

(4) 由最后一个已知样本时点的状态及所确定的状态转移概率矩阵，就可以得出未来的转移状态。在判定组合模型中各单项预测模型的预测精度状态时，若转移概率矩阵中各个单项预测模型的状态不同，则直接按照状态定义排序即可；如果状态相同，但是转移概率不同，则转移概率大的预测精度高；若状态相同，转移概率也相同，则按照各单项预测模型相同状态的近几期内的预测精度的均值进行排序。

(5) 由上述步骤得到各单项模型预测精度的高低顺序，求解得到组合模型的权系数，从而进行预测。

9.4.4 示例分析

某型机载导弹装备自服役以来，完成多次演习和战备任务，已成为海军主要的作战武器之一。在对海军航空兵场站弹药大队导弹保障情况进行调研分析的基础上，将备件需求主要影响因素归纳为固有故障率 λ （导弹定检故障率）、通电检测次数 X 、战备值班时间 T 、恶劣工作天气（室外温度高于 30°C 或低于 -5°C 、相对湿度高于 70% 等）占全年工作日的比例 Q 。该型导弹装备 1999—2011 年的装备保障无量纲化处理后的相关数据及某种常用备件的需求情况如表 9-9 所列。

1. 预测模型建立与评价

将 1999—2008 年的数据作为建模数据，构成样本空间；2009—2011 年的数据作为检验数据，构成预测区间。根据备件需求时间序列的非线性特征，选用三次指数平滑法进行预测，运用相对误差最小二乘法求得平滑系数 $\alpha = 0.63$ ；对 GM(1,1) 进行检验，其后验差比值 $C = 0.032 < 0.35$ ，小误差概率 $P = 1 > 0.95$ ，预测模型等级为优；运用网格搜索法对 WLS-SVM 的惩罚系数 C 和核参数 g 进行

优化, 当 $C = 1024$, $g = 0.0315$ 时均方根误差 (MSE) 达到最小值, 此时的 (C, g) 取值为最优值。选择误差指标预测误差平方和 (SSE)、平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE) 和均方百分比误差 (MSPE) 作为衡量精度的评价指标体系, 各种方法的预测结果如表 9-9 所列。

表 9-9 保障数据

年份	影响因素量化值				备件需求量/个
	λ /%	X /次	T /h	Q /%	
1999	0.0972	0.7176	0.0000	0.3869	20
2000	0.0716	0.5017	0.0000	0.3573	16
2001	0.0649	0.0000	0.1419	0.3247	14
2002	0.0715	0.5491	0.1581	0.3452	17
2003	0.0563	0.6224	0.3863	0.3628	22
2004	0.0238	0.7942	0.6467	0.2936	29
2005	0.0463	0.7168	0.4952	0.2594	23
2006	0.0321	0.6667	0.6943	0.2741	26
2007	0.0437	0.8563	0.7294	0.2150	31
2008	0.0575	1.0000	0.8867	0.3736	39
2009	0.0763	0.8934	0.7646	0.2383	34
2010	0.0981	0.7881	0.8358	0.3192	35
2011	0.1028	0.9415	1.0000	0.2875	42

由表 9-10 中各单项预测方法的预测值和预测精度, 根据式 (9-80) 可计算出 IOWHA 组合预测值, 将其代入式 (9-81) ~ 式 (9-85), 经整理化简得如下最优化模型:

$$\min \tau(w_1, w_2, w_3) = \left(\begin{matrix} 0.0096w_1^2 + 0.0183w_2^2 + 0.0631w_3^2 + \\ 0.02629w_1w_2 + 0.0441w_2w_3 + 0.0594w_1w_3 \end{matrix} \right)^{1/2}$$

$$\text{s.t.} \quad \begin{cases} w_1 + w_2 + w_3 = 1 \\ w_1 \geq 0, w_2 \geq 0, w_3 \geq 0 \end{cases}$$

将建模数据作为组合模型输入, 分别利用 GSA 和 IMGSA 算法对建立的组合预测模型进行求解, 其中初始种群数量取 40, 最大迭代次数取 80, 其算法寻优过程如图 9-11 所示。可见, IMGSA 在迭代次数较少的情况下就能达到最优, 而且适应度最为理想。

表 9-10 各种预测模型拟合结果及其评价

年份	实际 需求量	平滑指数法		GM(1,1)模型		WLS-SVM		改进 IOWHA	IOWHA	最优
		预测值	精度	预测值	精度	预测值	精度	组合模型	组合模型	加权法
1999	20	18.2037	0.9102	20.0000	1.0000	19.3608	0.9680	19.6901	20.8453	21.2247
2000	16	14.6561	0.9160	17.7253	0.9027	17.4634	0.9162	16.3149	17.2595	17.3489
2001	14	14.2109	0.9852	13.2702	0.9479	14.5660	0.9611	14.3172	15.3577	13.6684
2002	17	16.4062	0.9651	15.6366	0.9198	16.6547	0.9797	16.4783	18.1937	16.5925
2003	22	19.6153	0.8916	19.1795	0.8718	20.3741	0.9261	20.9089	21.2571	19.4876
2004	29	25.7968	0.8895	25.0468	0.8637	27.2583	0.9399	27.8634	28.4159	27.3261
2005	23	24.8746	0.9246	22.3574	0.9721	24.4169	0.9420	23.1365	24.3862	25.8462
2006	26	27.3827	0.9495	24.9861	0.9610	26.9572	0.9645	26.2538	26.5738	27.2798
2007	31	29.6869	0.9576	30.4937	0.9837	32.1526	0.9642	31.3912	30.5024	30.2953
2008	39	36.3594	0.9323	35.8415	0.9190	37.1563	0.9527	38.0063	37.8506	35.6486
SSE		35.5002		40.6245		16.3183		4.2689	9.9479	34.1830
MSE		0.0596		0.6374		0.4040		0.2066	0.3154	0.5847
MAE		1.6744		1.5914		1.1752		0.5466	0.9435	1.5681
MAPE		0.1297		0.1369		0.0601		0.0231	0.0466	0.1444
MSPE		0.0237		0.0255		0.0171		0.0072	0.0169	0.0209

从表 9-10 来看，基于改进 Theil 不等系数的 IOWHA 组合模型的评价指标值明显低于其他方法的评价指标值，而且其拟合精度与单项预测方法和传统的最优加权法相比也有一定程度的提高。另外，分别计算三种单项预测方法的预测值序列和基于改进 Theil 不等系数的 IOWHA 组合模型得到的预测值序列与实际值序列之间的 Theil 不等系数，得到 $\tau_1 = 0.0935, \tau_2 = 0.1843, \tau_3 = 0.0526, \tau = 0.0261$ 。因为 $\tau < \min\{\tau_1, \tau_2, \tau_3\}$ ，所以，由定义 9.6 可知基于改进 Theil 不等系数的 IOWHA 组合模型为优性组合预测。

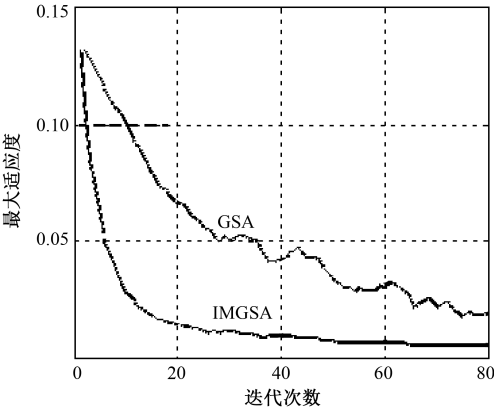


图 9-11 求解算法寻优过程曲线

2. 分析预测误差

根据表 9-10, 以预测精度为基准, 对其进行状态空间的划分, 各状态空间划分范围如下: 状态1 < 0.94, $0.94 \leq$ 状态2 ≤ 0.96 , 状态3 > 0.96, 如表 9-11 所列。

表 9-11 三种单项预测方法状态空间的划分

年份 状态	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
V_{1t}	1	1	3	3	1	1	1	2	2	1
V_{2t}	3	1	2	1	1	1	3	3	3	1
V_{3t}	3	1	3	3	1	1	2	3	3	2
注: V_{1t} 、 V_{2t} 、 V_{3t} 分别表示指数平滑法、GM(1,1)模型、WLS-SVM 所处的状态										

因为需要预测的 2009 年的预测精度未知, 所以 2008 年在计算时不参与。由表 9-11 可得三种单项预测模型各精度状态的元素个数, 如表 9-12 所列。结合表 9-11 和表 9-12 结果, 由各单项预测模型的初始状态概率及一步至三步状态概率转移矩阵可得到未来三期中各单项预测模型的预测精度状态, 如表 9-13 所列。由此可知, 2009 年各单项模型的预测精度顺序为 $x_{1t} > x_{3t} > x_{2t}$ 。同理, 2010 年和 2011 年各单项模型的预测精度顺序均为 $x_{3t} > x_{2t} > x_{1t}$, 进而可以确定组合模型的权系数。

表 9-12 各精度状态的元素个数

精度状态	元素个数		
	x_{1t}	x_{2t}	x_{3t}
状态 1	5	4	3
状态 2	2	1	1
状态 3	2	4	5
注: x_{1t} 、 x_{2t} 、 x_{3t} 分别表示指数平滑法、GM(1,1)模型、WLS-SVM			

表 9-13 各单项模型在预测期间的精度推测

模型	精度状态		
	2009	2010	2011
x_{1t}	3	1	1
x_{2t}	1	2	2
x_{3t}	2	3	3
注: x_{1t} 、 x_{2t} 、 x_{3t} 分别表示指数平滑法、GM(1,1)模型、WLS-SVM			

3. 模型应用与检验

由表 9-13 可知，组合预测模型为 $Y_{IOWHA} = x_{1t}$ 和 x_{3t} ，即在第一期预测中选择指数平滑法而在后两期的预测中选择 WLS-SVM 模型，能够更好地反映备件需求变化情况，并把它们的预测值作为备件需求量的比较基准值，进而计算出最终的预测结果，如表 9-14 和图 9-12 所示。

由表 9-14 可知，改进组合预测模型的预测性能好于各个单项预测方法及其他组合预测模型，而且性能稳定。另外，以 0.96 划定预测精度状态 2 和预测精度状态 3 的分界线与实际的预测精度相比较有一定的出入，但是预测精度的高低顺序是满足要求的。

表 9-14 不同方法的预测结果对比

年份	实际值	指数平滑法		GM(1,1)模型		LS-SVM		改进 IOWHA 组合		IOWHA 组合		最优加权法	
		预测值	RPE	预测值	RPE	预测值	RPE	预测值	RPE	预测值	RPE	预测值	RPE
2009	34	36.7329	0.0804	37.4531	0.1016	36.1865	0.0643	34.6972	0.0205	35.5895	0.0468	36.4839	0.0731
2010	35	34.2335	0.0219	36.4892	0.0425	37.3740	0.0678	35.7541	0.0215	36.1363	0.0325	38.3943	0.0970
2011	42	38.3587	0.0867	37.4527	0.1083	39.8592	0.0510	41.4634	0.0128	40.7269	0.0303	44.8517	0.0679
MRE		0.0630		0.0841		0.0610		0.0183		0.0365		0.0793	

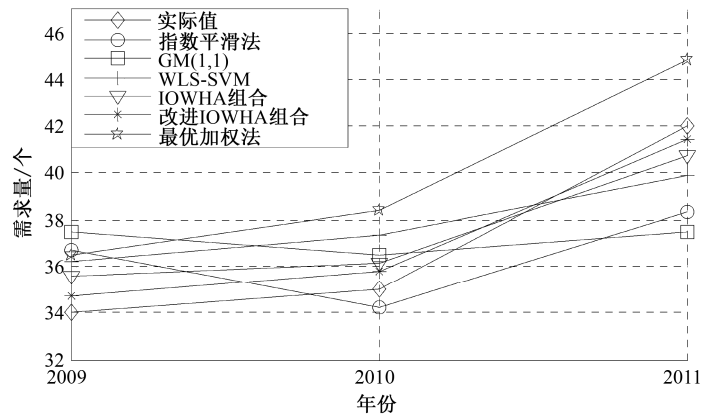


图 9-12 预测误差对比

参考文献

- [1] 赵建忠. 导弹装备备件需求预测与配置优化建模研究[D]. 烟台: 海军航空工程学院, 2013.
- [2] 赵建忠, 徐廷学, 刘勇, 等. 基于粗糙集、熵权以及改进支持向量机的导弹备件消耗预测[J]. 兵工学报, 2012, 33 (10): 1258-1265.
- [3] 赵建忠, 徐廷学, 葛先军, 等. 基于小波变换和改进 GM(1,1)-ARMA 的导弹备件消耗预测[J]. 北京航空航天大学学报, 2013, 39 (4): 553-558.
- [4] 赵建忠, 李海军, 叶文, 等. 改进系统备件满足率约束下的备件优化配置建模[J]. 兵工学报, 2013, 34 (9): 1187-1192.
- [5] 赵建忠, 徐廷学, 李海军, 等. 基于改进 Theil 不等系数的导弹备件消耗预测[J]. 系统工程与电子技术, 2013, 35 (8): 1681-1686.
- [6] 赵建忠, 李海军, 叶文, 等. 基于信息融合和改进 GUM(1,1)模型的故障预测[J]. 系统工程与电子技术, 2013, 35 (10): 2135-2140.
- [7] 赵建忠, 徐廷学, 李海军, 等. 基于小波分析的导弹装备备件需求组合预测[J]. 电子学报, 2013, 35 (10): 417-423.
- [8] 赵建忠, 叶文, 张磊. 基于数据融合和 IMUGM(1,1)模型的导弹装备故障预测[J]. 兵工学报, 2014, 35 (8): 1084-1089.
- [9] 赵建忠, 徐廷学, 叶文, 等. 基于数据融合和改进 MUGM(1, m,w)的导弹装备故障预测[J]. 系统工程与电子技术, 2015, 37 (4): 832-838.
- [10] 赵建忠, 欧阳中辉, 张磊, 等. 基于多智能体遗传算法优化的航空电子装备状态组合预测[J]. 兵工学报, 2016, 37 (4): 727-734.
- [11] 赵建忠, 邓建球, 叶文, 等. Combined forecast method of HMM and LS-SVM about electronics state based on MAGA[J]. 系统工程与电子技术, 2016, 27 (3): 730-736.
- [12] 赵建忠, 叶文, 刘勇, 等. Combined forecast method of electronics equipments based on improved HMM and LS-SVM[C]. Shenyang: Logistics Engineering Management and Computer Science (LEMCS2015): 154-157.
- [13] 赵建忠, 徐廷学, 郭宏超, 等. 基于改进最小二乘支持向量机的导弹备件消耗预测[J]. 贵州师范大学学报, 2012, 30 (2): 95-102.

- [14] 赵建忠, 徐廷学. 基于改进 GM(1,1)模型的导弹备件消耗预测[J].装备环境工程, 2012, 9 (3): 48-70.
- [15] 赵建忠, 徐廷学, 尹延涛, 等. 基于粗糙集和 BP 神经网络的导弹备件消耗预测[J].兵工自动化, 2012, 31 (7): 66-71.
- [16] 赵建忠, 徐廷学. 基于故障规律的导弹备件消耗预测[J].战术导弹技术, 2012 (5): 44-49.
- [17] 赵建忠, 徐廷学, 甄伟, 等. 基于改进支持向量机的导弹备件消耗预测[J].现代防御技术, 2013, 41 (2): 177-211.
- [18] 赵建忠, 徐廷学. 基于最大熵和仿真的间断工作导弹装备备件需求预测[J].现代防御技术, 2014, 42 (8): 87-94.
- [19] 徐廷学, 赵建忠, 李海军, 等. 基于 Logistic、Markov 和改进灰自助法的导弹备件需求预测[J]. 科技导报, 2013, 31 (16): 51-55.
- [20] 徐廷学, 赵建忠, 刘勇, 等. Fault forecasting of Missile equipment based on improved UGM(1,m,w) model[C].Shanghai: International Conference on Quality, Reliability,Risk,Maintenance,and Safey Engineering (QRRMSE2012): 746-749.
- [21] 董琪, 徐廷学, 赵建忠, 等. 基于调拨方案的两级库存备件配置优化研究[J]. 战术导弹技术, 2014, 165(3): 16-22.
- [22] 董琪, 徐廷学, 杨继坤, 等. 基于两级联合库存的低消耗备件配置优化研究[J]. 现代防御技术. 2015, 43 (2): 165-171.
- [23] 徐廷学, 董琪, 杨继坤, 等. 基于非泊松分布的不可修备件配置优化模型[J]. 科技导报, 2015, 33 (1): 95-99.
- [24] 张书君, 赵建忠. 基于改进灰色预测模型的军械维修器材消耗预测方法[J].四川兵工学报, 2011, 32 (5) .
- [25] 王相飞. 某型导弹备件消耗规律预测方法研究[D]. 烟台: 海军航空工程学院, 2010.
- [26] 金天球. 飞航导弹备件需求预测方法研究[D]. 烟台: 海军航空工程学院, 2009.
- [27] 董琪. 海军舰船特装器材配置优化与库存控制建模研究[D]. 烟台: 海军航空工程学院, 2015.
- [28] 徐宗昌. 装备保障性工程与管理[M]. 北京: 国防工业出版社, 2006.
- [29] 徐廷学. 装备综合保障工程理论与方法[M]. 北京: 兵器工业出版社, 2009.
- [30] 杨秉喜. GJB4355《备件供应规划要求》实施指南[M]. 北京: 总装备部技

- 术基础管理中心, 2006.
- [31] 韩兴才. 备件管理工程[M]. 北京: 蓝天出版社, 2003.
- [32] 孟祥辉, 徐宗昌, 刘文宝. 装备备件管理技术综述[J]. 装甲兵工程学院学报, 2009, 23 (4): 21-25.
- [33] 曲立. 备件需求预测与库存管理研究[D]. 北京: 北京科技大学, 2007.
- [34] 马应欣. 某型飞机航材定额管理研究[D]. 烟台: 海军航空工程学院, 2005.
- [35] 赵增庆. 空军场站航材库存优化分析[D]. 徐州: 徐州空军学院, 2004.
- [36] 翟承. 基于 ABC 分类法的航材库存管理研究[D]. 徐州: 徐州空军学院, 2004.
- [37] 邓聚龙. 灰理论基础[M]. 武汉: 华中科技大学出版社, 2002.
- [38] 刘思峰, 谢乃明. 灰色系统理论及其应用[M]. 北京: 科学出版社, 2008.
- [39] 肖新平, 毛树华. 灰色预测与决策方法[M]. 北京: 科学出版社, 2013.
- [40] 盖妹茜. 基于需求预测的备件库存协同管理研究[D]. 大连: 大连海事大学, 2007.
- [41] 陈华友. 组合预测方法有效性理论及其应用[M]. 北京: 科学出版社, 2008.
- [42] 唐小我. 现代组合预测和组合投资决策方法及应用[M]. 北京: 科学出版社, 2003.
- [43] 杨叔子, 吴雅. 时间序列分析的工程应用[M]. 武汉: 华中理工大学出版社, 1996.
- [44] 许绍杰. 基于需求预测的雷达器材保障系统研究[D]. 武汉: 空军雷达学院, 2011.
- [45] 李安贵, 张志宏, 孟艳, 等. 模糊数学及其应用[M]. 北京: 冶金工业出版社, 2005.
- [46] 夏乐天. 应用概率统计[M]. 北京: 机械工业出版社, 2008.
- [47] 飞思科技产品研发中心. 小波分析理论与实现[M]. 北京: 电子工业出版社, 2005.
- [48] 茆诗松, 程依明, 濮晓龙. 概率论与数量统计教程[M]. 北京: 高等教育出版社, 2004.
- [49] 孟祥辉. 面向动用的装甲装备备件需求预测与优化配置建模研究[D]. 北京: 装甲兵工程学院, 2011.
- [50] Craig C, Sherbrooke. Optimal inventory modeling of systems: multi echelon techniques[M]. Massachusetts: Kluwer Academic Publishers Group, 2004.

- [51] 贺步杰. 装备备件最优库存建模——多级技术[M]. 2 版. 北京: 电子工业出版社, 2008.
- [52] 中国人民解放军总装备部. GJB 4050—2000 武器装备维修器材保障通用要求[M]. 北京: 总装备部军标出版发行部, 2000.
- [53] 刘晓芹. 电子装备预测与健康管理关键技术研究[D]. 石家庄: 军械工程学院, 2010.
- [54] 胡运权, 冯玉强, 孙文俊. 运筹学导论[M]. 8 版. 北京: 清华大学出版社, 2007.
- [55] Dasgupta D. Artificial immune systems and their application[M]. Berlin: Springer-Verlag, 1999.
- [56] Suykens J A K. Least squares support vector machines[M]. Singapore: World Scientific Publishing Co Pte Ltd, 2002.
- [57] Percivald B P, Walden A T. Wavelet methods for time series analysis[M]. England: Cambridge University press, 2000.
- [58] 董肇君, 等. 系统工程与运筹学[M]. 北京: 国防工业出版社, 2003.
- [59] 黄贯虹, 方刚. 系统工程方法与应用[M]. 广州: 暨南大学出版社, 2005.
- [60] 黄梯云. 智能决策支持系统[M]. 北京: 电子工业出版社, 2001.
- [61] 杨善林. 智能决策方法与智能决策支持系统[M]. 北京: 科学出版社, 2005.
- [62] 李业. 预测学[M]. 广州: 华南理工大学出版社, 1998.
- [63] 陈学楚, 张诤敏, 陈云翔, 等. 装备系统工程[M]. 2 版. 北京: 国防工业出版社, 2005.
- [64] 牟讴红. 某型机备件需求确定过程和方法的研究[D]. 北京: 北京航空航天大学, 2000.
- [65] 赵淑舫. 基于维修理论基础的航材需求预测方法研究[D]. 南京: 南京航空航天大学, 2001.
- [66] 伏洪勇. 航空备件计划模型与方法研究[D]. 北京: 北京航空航天大学, 2002.
- [67] 宁宣熙, 刘思峰. 管理预测与决策方法[M]. 北京: 科学出版社, 2003.
- [68] 王宏生. 人工智能及其应用[M]. 北京: 国防工业出版社, 2006.
- [69] 中国人民解放军总装备部. GJB 4355—2002, 备件供应规划要求[S]. 北京: 2003.
- [70] 裴鹿成. 蒙特卡罗方法及其应用[M]. 北京: 海洋出版社, 1998.
- [71] Dubi A. Monte Carlo applications in systems engineering[M]. New York: John Wiley&Sons Inc, 2007.

- [72] Evans R, Olson L. 数据、模型与决策[M]. 杜本峰, 译. 北京: 中国人民大学出版社, 2006.
- [73] 吴翊, 李永乐, 胡庆军. 应用数理统计[M]. 长沙: 国防科技大学出版社, 2008.
- [74] 国防科工委军用标准化中心. GJB 5361 飞机备件目录编制要求[S]. 北京: 国防科学技术工业委员会, 2005.
- [75] 国防科工委军用标准化中心. GJB 2528 舰船备品备件和供应品配置一般规定[S]. 北京: 国防科学技术工业委员会, 1995.
- [76] 张帅. 某型舰载机备件库存研究[D]. 烟台: 海军航空工程学院, 2015.
- [77] 周永务, 王圣东. 库存控制理论与方法[M]. 北京: 科学出版社, 2009.
- [78] 阮旻智. 多级维修供应模式下舰船装备备件的配置优化方法研究[D]. 武汉: 海军工程大学, 2012.
- [79] 任敏, 陈全庆, 沈震. 备件供应学[M]. 北京: 国防工业出版社, 2013.
- [80] Sherbrooke C C. 装备备件最优库存建模—多级技术[M]. 北京: 电子工业出版社, 2008.
- [81] Sheldon M. 应用随机过程[M]. 北京: 人民邮电出版社, 2011.
- [82] 蔡自兴, 徐光佑. 人工智能及其应用[M]. 北京: 清华大学出版社, 2004.
- [83] 徐国祥. 统计预测和决策[M]. 上海: 上海财经大学出版社, 1998.
- [84] 刑文训, 谢金星. 现代优化计算方法[M]. 北京: 清华大学出版社, 1999.
- [85] 冯尚友. 多目标决策理论方法与应用[M]. 武汉: 华中理工大学出版社, 1990.
- [86] Bonezek R H, Holsapple C W and Winston A B. Foundation of decision support systems[M]. New York: Academics press, 1998.
- [87] 杨善林. 智能决策方法与智能决策支持系统[M]. 北京: 科学出版社, 2005.
- [88] 张连武, 赵方庚, 陈学军. 美军装备备件需求预测应用现状及启示[J]. 兵工自动化, 2013, 32(7): 34-37.
- [89] 赵保粮. 供应链管理[M]. 北京: 对外经济贸易大学出版社, 2009.
- [90] 国防科工委军用标准化中心. GJB 2528 舰船备品备件和供应品配置一般规定[S]. 北京: 国防科学技术工业委员会, 1995.
- [91] 刘学锋, 王振安, 李海丰. 装备维修器材储备供应制约因素及对策分析[J]. 通用装备保障, 2011, 119(5): 40, 41.
- [92] 陈懿, 梁德重. 备品备件工作管理的问题及对策[J]. 装备制造技术, 2012 (11): 97-99.

- [93] Sherbrooke C C. Optimal inventory modeling of system: Multi-Echelon Techniques [M]. Springer. 2004.
- [94] Godoy R Pascual. Knights. Critical spare parts ordering decisions using conditional reliability and stochastic lead time [J]. Reliability Engineering & System Safety, 2013, 119(0): 199-206.
- [95] 朱绍强, 李寿安, 李为吉, 等. 具有最大可用度的航空备件供应模型[J]. 空军工程大学学报, 2005, 6(2): 22-24.
- [96] 陶小创, 郭霖瀚, 肖波平, 等. 基于备件保障概率分配的备件需求量预测模型[J]. 兵工学报, 2012, 33(8): 975-979.
- [97] De Smidt-Destombes S, Van Der Heijden C., Van Harten. Joint optimisation of spare part inventory, maintenance frequency and repair capacity for k-out-of-N systems [J]. International Journal of Production Economics, 2009, 118(1): 260-268.
- [98] De Smidt-Destombes S, Van Der Heijden C, Van Harten Aart, Availability of -out-of- systems under block replacement sharing limited spares and repair capacity [J]. International Journal of Production Economics, 2007, 107(2): 404-421.